Studies on the Genetic Diversity of *Haemophilus influenzae* type b

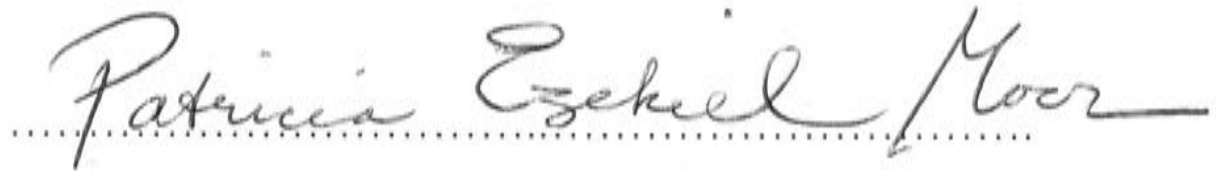
by

Patricia Ezekiel Moor

A thesis submitted in November 2002 for the degree
of Doctor of Philosophy
of the Australian National University

DECLARATION

"I declare that all the work described in this thesis, unless otherwise acknowledged, is the original research of the author and has not been submitted for a higher degree to any other university or institution."



Patricia Ezekiel Moor

ACKNOWLEDGMENTS

The work presented in this thesis was undertaken in the Division of Biochemistry and Molecular Biology in the Faculty of Science at the Australian National University with funds and resources provided by this Division and the Graduate School. An Australian Postgraduate Research Award provided the financial support that enabled me to undertake this work.

No achievement stands on the shoulders of one person and I owe thanks to many. For my husband Rick, I simply give thanks. To my son Adam, who never failed to add joy to my life during this adventure in science, I also give thanks. Yvonne Moor inspires my warmest gratitude for giving me the gift of time and understanding on so many occasions. My parents, Stanley and Ann Halas Ezekiel, who have always been there, are responsible for providing me with the education and upbringing that ultimately made this thesis possible. Thank you all for your love and support.

Professor Gwendolyn L. Gilbert deserves my thanks and admiration. She generously provided isolates and her input, suggestions and superb editing were invaluable.

I owe thanks to Dr Peter R. Stewart for giving me the opportunity to begin this project and for supporting me in the first months of its development. Thanks go to Dr Peter Collignon for providing access to so many *Haemophilus* isolates and for his interest and help in this endeavour. I am indebted to Dr Adrian Gibbs and John Armstrong of the Bioinformatics Facility for assistance with dendrograms and the RAPDistance Programs. After Dr Stewart's retirement, Professor Christopher Bryant took on the role of my supervisor until his own retirement and I thank him for his support.

Dr Ron Weir became my third primary supervisor for whom I am grateful. His faith in my efforts and his persistence to see me through my program made all the difference. His unfailing support and that of the Division and the Graduate School is appreciated.

My kindest thanks go to Professor Allan Cripps and Associate Professor Jennelle Kyd and members of the Canberra Centre for Mucosal Immunology in the Gadi Research Centre at the University of Canberra for providing laboratory support for important studies. Funds from the University Research Grants Scheme of the University of Canberra supported that work.

A few people in less formal roles who I met along the way should be mentioned for the help they gave me. Wafa El-Adhami earns my appreciation for putting up with me and for what I could learn from her. Barbara Inglis was a technical resource whose help kept the early experiments rolling. Bill Nicholson, head technical officer, and Neville Whitworth, electrician extraordinaire, were generous and helpful beyond all expectations when they were needed most. Huan Pham, seeking her PhD "on the outside", introduced me to a rare kind of strength of the human spirit that inspires me still.

Finally, my sincerest thanks go to Jan Bell and the other generous members of the scientific community who sent me isolates.

ABSTRACT

Haemophilus influenzae type b (Hib) is a major cause of serious invasive disease in children worldwide. Before the introduction of a vaccine in 1992 in Australia, it was the commonest cause of serious bacterial infections in children under the age of five years and the incidence of Hib disease was reported to be the highest in the world in Aboriginal populations. It is not clear what characteristics, if any, of the bacterium are associated with its epidemiology and little is known about its diversity in Australia. This thesis describes a study that used pulsed field gel electrophoresis (PFGE) to analyse the overall molecular genetic characteristics of 213 pre-vaccine Aboriginal and non-Aboriginal isolates of Hib from varied rural and urban regions in Australia and 20 post-vaccine isolates from an Aboriginal community. The objectives were to examine the genetic diversity of Hib to determine whether distinctive types are associated with its epidemiology and virulence and to generate baseline data that would permit assessment of the effects of vaccination on the genetic diversity and population structure of Hib.

Among the pre-vaccine isolates 69 different types of *Sma*I-RFLP patterns with 17 or fewer fragments were well resolved by PFGE. Closely related and more distantly related genomic types were easily discriminated. Two types, designated F2a and A8b, accounted for 43% of the sample and include almost one-half of the Aboriginal isolates and over one-third of the non-Aboriginal isolates, respectively. The types fell into 7 major clusters in a tree-like dendrogram that is consistent with a clonal population structure, as is the presence of two predominant types. The results reveal a clonal population structure and extend the geographical range over which the clonality of Hib has been demonstrated.

Two highly divergent groups comprising 7 types that represent 50 (66%) Aboriginal isolates and 6 (4%) non-Aboriginal isolates and another genetically distinct group comprised of 39 types representing 7 (9%) Aboriginal isolates and 106 (77%) non-Aboriginal isolates revealed a deep genetic divergence between the majority of Aboriginal and non-Aboriginal isolates. Overall, relatively fewer and more diverse types were found among the Aboriginal isolates compared to those from non-Aboriginals. A single clone predominated in each of the urban and rural areas studied except one where two predominant clones were identified.

The genetic distance separating the important major lineages ranged from a similarity of 20% to 50% yet there is apparently an equivalent ability to cause disease among them. No association between RFLP type and age, sex, date of collection, or disease manifestation was found. Because of the lack of Aboriginal isolates from urban areas and non-Aboriginal isolates from rural areas the association, if any, between geographic location and type could not be determined. However, there is a strong association (chi square, $p>0.001$) between the origin of isolates from Aboriginal children and RFLP type F2a and the origin of isolates from non-Aboriginal children and RFLP type A8b.

Important information about the dynamics of Hib strain diversity in the pre- and post-vaccine periods was revealed. Pre-vaccine PFGE types were found to be circulating several years after the introduction of vaccines in a remote Aboriginal community. Of 8 *Sma*I-RFLP types found among 20 post-vaccine isolates 5 are indistinguishable from types found among the pre-vaccine isolates. The same type (F2a) predominates in both the pre- and post-vaccine samples. The three post-vaccine types not found in the pre-vaccine sample are closely related to two of the pre-vaccine types. This divergence may suggest a shift in population structure but because of the relatively low number of isolates studied it cannot be ruled out that less frequent types were not recovered in the pre-vaccine sample.

The validity of the PFGE results was substantiated by several comparative analyses. Enterobacterial repetitive intergenic consensus (sequences)-PCR RFLP analysis and the results of *cap* locus, P2 and OMP26 hybridisation studies did not contradict the overall genetic relationships revealed among Hib isolates using PFGE. Differences found between Aboriginal and non-Aboriginal Hib types using PFGE demonstrate its practical discriminatory power for analysing large numbers of Hib isolates. Analysis of the huge similarity matrices required for large samples has been facilitated by the availability of computer programs and, with standardisation of PFGE methods, the database developed here will allow inter- and intra-laboratory comparison of results. This will improve surveillance by increasing the capacity to detect new strains of Hib and monitor the presence of old strains. This methodology will be useful in monitoring future Hib isolates from children in whom the vaccine has failed, unimmunised children, and adults.

ABSTRACTS AND PUBLICATIONS

ARISING FROM THIS THESIS

2002 "*Haemophilus influenzae* type b (Hib) genotypes in an Aboriginal population before and after the introduction of Hib conjugate vaccines". PE Moor and HC Smith-Vaughan (in preparation)

2001 Courtesy picture provided of restriction fragments from pulsed field gel electrophoresis of Hib genomic DNA digested with *Sma*I. "Microbiology: a human perspective", third edition, p. 257. Editors EW Nestor, DG Anderson, CE Roberts, NN Pearsall, and MT Nester.

1999 "Pulsed field gel electrophoresis used to investigate the genetic diversity of *Haemophilus influenzae* type b isolates in Australia shows differences between Aboriginal and non-Aboriginal isolates". PE Moor, PC Collignon, GL Gilbert. Journal of Clinical Microbiology, May 1999, 37(5): 1524-1531.

1998 "PFGE analysis of RFLPs shows differences between Hib strains isolated from Aboriginal and non-Aboriginal children in Australia". American Society for Microbiology Annual General Meeting, Atlanta, Georgia. PE Moor, GL Gilbert, PC Collignon. Abstract C-65, p. 142, poster presentation.

1997 "The genetic diversity of noncapsulate *Haemophilus influenzae*". P Ezekiel Moor. International Conference on Acute Respiratory Infections, Canberra, Australian Capital Territory. Abstract W4B, p. 96, proffered paper.

1995 "The identification of invasive *Haemophilus influenzae* using molecular techniques". P Ezekiel Moor and R Kelly. Australian Society for Microbiology and the New Zealand Microbiological Society Combined Scientific Meeting and Exhibition, Canberra, Australian Capital Territory. Abstract P16.5, proffered paper.

1994 "The application of restriction fragment length polymorphism analysis using pulsed field gel electrophoresis to the epidemiology of *Haemophilus influenzae* type b". P Ezekiel Moor. Australian Society for Microbiology and the New Zealand Microbiological Society Combined Scientific Meeting and Exhibition, Melbourne, Victoria. Abstract GP14, p. A-112, poster presentation.

1993 "The application of restriction fragment length polymorphism analysis using pulsed field gel electrophoresis to the epidemiology of *Haemophilus influenzae* type b". P Ezekiel Moor. Australian Society for Microbiology and the New Zealand Microbiological Society Combined Scientific Meeting and Exhibition, Perth, Western Australia. Abstract GL11, p. A-103, poster presentation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

A	Aboriginal
ACT	Australian Capital Territory
AP-PCR	arbitrarily primed-polymerase chain reaction
bp	basepair
BSA	bovine serum albumin
CHEF	contour clamped homogeneous electric field
D	Simpson's index of diversity
dNTP	deoxynucleotide 5´- triphosphate
DNA	deoxyribonucleic acid
DOC	date of collection
EDTA	ethylenediaminetetra-acetic acid
ERIC	enterobacterial repetitive intergenic consensus (sequences)
ESP	EDTA/ N-lauroyl sarcosine/proteinase K
ET	electrophoretic type
HbOC	polyribosylribitol phosphate conjugated to non-toxic oligosaccharide mutant diphtheria toxin (vaccine)
Hia	*Haemophilus influenzae* adhesion protein (autotransporter)
Hap	*Haemophilus* adhesion and penetration protein (autotransporter)
Hib	*Haemophilus influenzae* type b
kb	kilobase (1,000 base pairs)
kDa	kilodalton
λ	Lambda phage DNA
λ-H	Lambda phage DNA digested by *Hin*dIII
LOS	lipooligosaccharide
Mb	megabase (1,000,000 base pairs)
µm	micron or micrometer
min	minute(s)
MLEE	multilocus enzyme electrophoresis
MSHR	Menzies School of Health Research
MRSA	methicillin resistant *Staphylococcus aureus*
n-A	non-Aboriginal
na	not applicable
NCHi	non-capsulate or non-encapsulated *Haemophilus influenzae* (technically, NCHi refers to *H. influenzae* strains that do not possess the *cap* loci (and, thus, do not demonstrate the presence of capsular material by serotyping) and are otherwise genetically dissimilar to encapsulated strains; in this thesis NCHi is used to designate *H. influenzae* isolates in which the absence of the *cap* locus has been determined, though, in popular practice this designation is still used interchangeably with NTHi)
NJTREE	Neighbor Joining Tree program
NTHi	non-typeable *Haemophilus influenzae* (technically, NTHi refers to *H. influenzae* strains that do not demonstrate the presence of capsular material by serotyping and which have not been analysed by molecular methods for the presence of the *cap* locus; this is the traditional designation for non-encapsulated strains and is used in this thesis when the absence of the *cap* locus has not been determined, though, in popular practice this designation is still used interchangeably with NCHi)
NT	Northern Territory
OMP	outer membrane protein
PCR	polymerase chain reaction
PE	phosphatidylethanolamine
PEM	Patricia Ezekiel Moor
PRP	polyribosylribitol phosphate

PRP-D	polyribosylribitol phosphate conjugated to diphtheria toxoid (vaccine)
PRP-OMP	polyribosylribitol phosphate conjugated to an outer membrane protein complex of *Neisseria meningitidis* (vaccine)
PRP-T	polyribosylribitol phosphate conjugated to tetanus toxoid (vaccine)
PFGE	pulsed field gel electrophoresis
RAPD	randomly amplified polymorphic DNA
RFLP	restriction fragment length polymorphism
RT	room temperature
sBHI	supplemented brain heart infusion
SDS	sodium dodecyl sulfate
TAE	Tris-acetate-EDTA buffer
TBE	Tris-borate-EDTA buffer
TDRAW	Table drawing program
TE	10 mM Tris, 1 mM EDTA, pH 8.0
TX	Texas
WCP	whole cell polypeptide
WHO	World Health Organisation

Thus have I made as it were a small
globe of the intellectual world, as truly
and faithfully as I could discover.

FRANCIS BACON (1605)

CHAPTER 1

Introduction

1.1 Introduction

Haemophilus influenzae type b (Hib) is a bacterium that causes life threatening invasive disease in children, particularly among those under 5 years of age. Before the establishment of widespread Hib vaccination in Australia its epidemiology was characterised by an extremely high incidence of invasive disease among Aboriginal children compared to non-Aboriginal children (Hanna and Wild, 1991; Hanna, 1992; Hanna, 1995), a lack of epiglottitis among Aboriginals (Hanna, 1992), and significant differences in the incidence of epiglottitis between urban areas (Gilbert *et al.*, 1990; McIntyre *et al.*, 1991; Hanna *et al.*, 1992; McGregor *et al.*, 1992). The reasons for the unique epidemiology of Hib remain largely unknown and there is little information about its population structure and genetic diversity in Australia.

The project presented in this thesis was initiated to assess the genetic diversity of isolates of Hib in order to answer questions concerning its epidemiology and population structure in Australia. The advent of the technology that allows us to examine bacterial genetic diversity at a molecular level occurred during the same time as the development of the Hib conjugate vaccines. For the first time ever, the opportunity exists to monitor the effects of immunization by examining, not only, the change in the incidence of disease, but also, the change, if any, in the molecular genetic diversity of the population structure of the target microorganism. The overall purpose of this study has been to analyse the molecular genetic characteristics of clinical isolates of Hib to determine whether distinctive types are associated with its epidemiology and to develop a typing system for monitoring *H. influenzae* isolates from children in whom the vaccine has failed, unimmunised children, and adults. The introduction that follows presents the general characteristics of Hib, an overview of typing methods used to study Hib and a review of our current understanding of the population structure and genetic diversity of *H. influenzae*. It concludes with a description of the aims of the study.

1.2 Haemophilus influenzae: Historical Portrait of a Bacterium

The bacterium known as *H. influenzae* was first seen by Robert Koch in 1883 when he found small Gram-negative bacilli in the pus of patients with conjunctivitis. It was separately discovered and initially cultured from similar material in 1886 by John Weeks and came to be known as the Koch-Weeks bacillus and then *H. aegyptius*. It was

only much later in the 1970s that we learned from Mogens Killian to include *H. aegyptius* in a biotype of *H. influenzae* (Turk, 1982a). The name *H. influenzae* is based on Rickard Pfeiffer's erroneous suggestion during the influenza pandemic of the 1890s that the 'influenza bacillus' he cultured from the sputa and lungs of patients was the cause of the disease (Pfeiffer, 1893). The American Association of Bacteriologists proposed the name *Haemophilus influenzae* for Pfeiffer's bacillus in 1920 long before the viral origin of influenza was determined in 1933. By this time the variability in the pathogenicity of *H. influenzae* was recognised and Margaret Pittman had established the significance of a capsule in the increased pathogenicity of cerebrospinal fluid isolates (Pittman, 1931). While it seems likely that in at least some outbreaks of influenza there may be a virus-*Haemophilus* synergism, aside from Pfeiffer's association of it with the pandemic of the late 1800s and recognition of its presence in the 1918 pandemic (apparently as a secondary invader), *H. influenzae* has done little since to earn its species name (Turk, 1982b; Willett, 1988).

The true picture of the range of pathogenic activities of *H. influenzae* became clearer in the period 1930-1960 (Turk, 1984) and knowledge of its epidemiology, pathogenicity, and genetics grew dramatically in the 25 years before and since the 1993 centenary of its cultivation by Pfeiffer. One of the important early observations was the description of the relationship between a high incidence of Hib meningitis and the low bactericidal power of serum in children 2 months to 3 years of age (Fothergill and Wright, 1933). The recognition that this reflected a lack of serum antibody capable of providing protection against invasive disease led to the development of a first generation vaccine consisting of the purified capsular polysaccharide of Hib, polyribosylribitol phosphate. This vaccine proved to be successful in older children but not in those under 2 years of age, the age group with the highest burden of disease. The conjugation of capsular polysaccharide to protein was a vital leap in technology that resulted in Hib vaccines capable of protecting infants (Heath, 1998). The effectiveness of the conjugate vaccines is such that in most areas where they have been introduced the incidence of Hib disease has been significantly reduced and its clinical significance diminished. Hib continues, however, to be a serious pathogen in populations where the vaccine is not available and the challenge remains to implement worldwide vaccination in areas where epidemiological studies warrant its use.

In early 1995 the determination of the complete sequence of a bacterial genome was accomplished for the first time ever when the final piece of sequence data was slotted into the circular chromosome of *Haemophilus influenzae* Rd KW20 (Fleischmann *et al.*, 1995). Little known for being the strain from which one of the first DNA restriction enzymes were purified (Smith and Wilcox, 1970), this further honour was bestowed on *H. influenzae* Rd KW20 largely due to the its relatively small size (1.8 Mb).

Remarkable achievements in the last two decades have deeply changed our understanding of the epidemiology, clinical significance and laboratory research of *H. influenzae*. The dramatic effect of the conjugate vaccines in preventing invasive disease and the publication of the *H. influenzae* genome sequence are just two of the extraordinary scientific achievements that, not only, shed light on our complex and long standing relationship with *H. influenzae*, but, illuminate our vision of the microbial world in general.

1.3 The Genus, Haemophilus

Haemophilus organisms are among the smallest of bacteria. The genus is restricted to a group of bacteria <1 μm in width and 1-1.5 μm in length. They are chemoheterotrophic, non-motile, non-spore forming aerobic or facultatively anaerobic, oxidase-positive (rarely negative), catalase-positive (rarely negative), Gram-negative rods that sometimes form threads and show marked pleomorphism (primarily when recovered from cerebrospinal fluid or broth cultures) but they are most often found in a uniform rod shape. The curved ends of the short bacilli makes many appear round and, hence, they are referred to as coccobacilli. Almost all haemophili require preformed growth factors present in blood, particularly X and V factor (Killian and Biberstein, 1984) as shown in Table 1.1. X factor is actually a group of heat stable compounds provided by several iron-containing pigments and V factor is nicotinamide adenine dinucleotide or nicotinamide adenine dinucleotide phosphate. The very name of the genus comes from the requirement for these accessory growth factors found in blood and is derived from the Greek words meaning "blood loving". The genus is currently classified in the family Pasteurellaceae that includes members of the genera Pasteurella and Actinobacillus (Killian and Biberstein, 1984).

Haemophili occur as obligate parasites on the mucous membranes of both humans and animals and, with the possible exception of *H. parainfluenzae*, are host specific (Killian and Biberstein, 1984). The ten human species and six animal species currently listed in the genus in Bergey's Manual of Determinative Bacteriology (Holt *et al.*, 1994) are shown in Table 1.1. The human species are found almost exclusively as members of the normal microbial flora of the oropharynx and they possess varying propensities for causing disease. They are also found, but much less frequently, in the normal flora of the genital tract with the exception of *H. ducreyi* which is a primary pathogen of the genital tract.

Table 1. 1 Growth factor requirements of *Haemophilus* species isolated from humans[a]

Species	Factor Requirement	
	X	V
H. influenzae[b]	+	+
H. aegyptius[c]	+	+
H. parainfluenzae	–	+
H. haemolyticus	+	+
H. parahaemolyticus	–	+
H. aphrophilus	–	–
H. paraphrophilus	–	+
H. paraphrohaemolyticus	–	+
H. segnis	–	+
H. ducreyi[d]	+	–

Data obtained from Kilian, M. and Bilberstein, E.L. in Bergey's Manual of Systematic Bacteriology, 1984.
[a] animal species include *H. pleuropneumoniae, H. haemoglobinophilus, H. parasuis, H. paragallinarium, H. paracuniculus, H. avium* (changes in classification are found in Bergeys' Manual of Determinative Bacteriology, 9th edition, 1994)
[b] includes 6 encapsulated types, and non-typeable *H. influenzae*
[c] now classified as *H. influenzae* biogroup *aegyptius*
[d] evidence is accumulating that *H. ducreyi* may be only distantly related to the genus (Dewhirst *et al.*, 1992)

1.4 The Species, Haemophilus influenzae

Haemophilus influenzae is the most important pathogen among the species that colonise the human host. With the occasional exception of the human genital tract, the human nasopharynx is the sole ecological niche occupied by this organism. It is not found in the environment and does not colonise or cause disease in other animal species.

It may make one of six chemically distinct capsular polysaccharides designated types a-f, but among the six capsule producing strains of *Haemophilus influenzae* only type b

isolates commonly cause systemic disease. Those strains of *H. influenzae* that do not produce a capsule and are serologically non-reactive in capsular antisera are referred to as non-encapsulated or non-typeable *H. influenzae* (NCHi or NTHi). The NCHi include a subset of a genetically distinct group described as a cryptic genospecies that is distantly related to *H. influenzae* and *H. haemolyticus* (Quentin *et al.*, 1993).

Strains of *H. influenzae* can be separated into biotypes (biovars) using a battery of biochemical tests. Kilian's system of biotyping based on indole production and ornithine decarboxylase and urease activities (Kilian, 1976) is currently used to divide *H. influenzae* into biotypes I-VIII as shown in Table 1.2. The usefulness of biotyping for epidemiological purposes or for studying population structure is limited but valuable.

It has been reported recently that the presence of active copper- and zinc-containing superoxide dismutase in isolates of the cryptic genospecies (non-capsulate strains of biotype IV and to a lesser extent biotype II), responsible for urogenital, neonatal, and mother-infant infections, can be used as a biochemical marker to discriminate them from *H. influenzae* sensu stricto strains (Langford *et al.*, 2002).

Table 1. 2 Biotypes of *Haemophilus influenzae*[a]

Biotype	Indole Production	Ornithine Decarboxylase Activity	Urease Activity
I[b]	+	+	+
II	+	–	+
III	–	–	+
IV	–	+	+
V	+	+	–
VI	–	+	–
VII	+	–	–
VIII	–	–	–

Modified from Kilian, M. 1991. *Haemophilus. In* Balows, A., *et al.* editors: *Manual of clinical microbiology*, 5th edition, American Society for Microbiology.
[a]found among both encapsulated and non-encapsulated isolates
[b]most Hib isolates fall into this biotype

1.5 Pathogenesis and Clinical Significance of Hib

Prior to the introduction of effective vaccines, disease caused by Hib was a worldwide problem of major importance in both developing and developed countries (Turk and May, 1967; Turk, 1982a). In areas where the vaccine has been introduced, primarily in developed countries, Hib disease has been virtually eradicated (Peltola *et al.*, 1992; Adams *et al.*, 1993). However, in those areas of the world, such as, in Asia and Africa, where Hib vaccines are not routinely used, Hib disease is still a significant cause of childhood morbidity and mortality, with devastating effects on the most vulnerable of all, babies and small children (Neuzil, 2000). Currently, less then 2% of Hib diseases worldwide are prevented because some 80% of the 215 countries reporting to the World Health Organisation (WHO) are not using Hib vaccines in their national immunisation programs (Peltola, 2000). Subsequently, some 175 countries and 118 million children are without protection against life-threatening Hib infections. According to the WHO, Hib still causes 3 million cases of serious disease and 400,000 to 700,000 deaths per year in young children worldwide (Peltola, 2000).

Implementation of universal coverage of Hib vaccines will require overcoming obstacles of titanic proportions for which the outlook is currently quite gloomy. These include perceptions of low disease burden due to underestimates of Hib disease, limited availability and the high cost of vaccines, and motivation for vaccine programs (Neuzil, 2000). Thus, despite the widespread use of Hib vaccines making the goal to eliminate Hib disease feasible (Heath, 1998; Campos, 2001), it remains a major global infectious disease challenge for the foreseeable future.

1.5.1 Clinical manifestations of Hib disease

A wide variety of *Haemophilus* infections may occur in humans, ranging from infections that are frequent and easily managed (e.g., conjunctivitis and otitis media) to those that are potentially life threatening and less frequent or rare (e.g., meningitis, pericarditis, and Brazilian purpuric fever) (Campos, 1995). *Haemophilus influenzae* is the species that most frequently causes infection and invasive diseases are usually caused by Hib, the *H. influenzae* strains possessing the type b capsule. In most industrialised countries in the pre-vaccine era, Hib was a frequent cause of invasive diseases in previously healthy children under the age of 5 and was the most common

cause of meningitis in children between 1 month and 3 years of age. This disease pattern is linked to the time after maternal antibody is lost and before the developing immune system is capable of protecting the infant. The other encapsulated strains (types a, c, d, e, and f) and non-encapsulated *H. influenzae* may also cause invasive disease but Hib has been shown to cause approximately 90% of the systemic *Haemophilus* infections in industrialised societies (Jordens and Slack, 1995).

While meningitis is the most common manifestation of invasive Hib disease, worldwide studies in the 1970s to 1990s showed that over 90% of Hib infections manifested as six "classical" entities: meningitis, epiglottitis, bacteremic pneumonia, septicemia, cellulitis, and osteoarticular infections (more often septic arthritis than osteomyelitis) (Peltola, 2000). Other clinical manifestations constituted only 3% of the total number, whereas multifocal cases were diagnosed in 6%. Table 1.3 shows the distribution of Hib diseases taken from data for 3,391 patients in 21 studies from various parts of the world.

Table 1. 3 Worldwide spectrum of all classical Hib diseases[a]

Diagnosis	%
Meningitis	52
Pneumonia[b]	12
Epiglottitis	10
Septicemia	8
Cellulitis	5
Osteoarticular	4
Others	3
Multifocal	6

Table modified from Peltola, 2000.
[a]data obtained in the 1970s-90s for 3,391 patients in 21 studies from various parts of the world
[b]non-bacteremic pneumonia excluded

A clinical review of Hib disease in the pre-vaccine era in Melbourne, Victoria in Australia showed that three-quarters of the patients presented with either meningitis or epiglottitis (Gilbert *et al.*, 1995). It is of interest that epiglottitis accounts for 40% of invasive Hib disease in Victoria and that this is twice that found in Sydney (McIntyre *et al.*, 1991) or Western Australia (Hanna *et al.*, 1992) and is four times that found worldwide (Peltola, 2000). This difference in incidence, as well as, age related issues are discussed in the section of this chapter on epidemiology. Except for epiglottitis, the distribution of the classical Hib disease manifestations was roughly the same wherever it was studied (Peltola, 2000). In Victoria, aside from the increased incidence of

epiglottitis with a concomitant reduction in Hib meningitis, the distribution of Hib disease was similar to the worldwide distribution. The distribution of primary diagnoses in 235 patients in Victoria is shown in Table 1.4.

Table 1. 4 Primary diagnoses in 235 children from Victoria, Australia[a] with invasive Hib disease showing an increased incidence of epiglottitis and a concomitant reduction in meningitis as compared to the worldwide distribution of Hib disease

Diagnosis	*n*	%
Epiglottitis	94	40
Meningitis	84	36
Pneumonia	20	9
Cellulitis	16	7
Bacteremia	14	6
Arthritis/osteomyelitis[b]	6	3
Pericarditis	1	0.4

Table modified from Gilbert, 1995.
[a]Patients admitted to the Royal Children's Hospital, Melbourne, Victoria between 1 February 1988 and 31 January 1990 were included in the study.
[b]Two additional patients who presented with meningitis developed arthritis after admission.

Specificity of certain bacteria for certain species, even certain tissues, and even certain cells is one of the most amazing of all biological phenomena. For example, the polio virus kills only the anterior horn cells of the spinal cord of humans (and occasionally other primates). Similarly, the ability of Hib to infect the epiglottis is characteristic as epiglottitis in children is rarely attributed to other microorganisms (Hanna *et al.*, 1992). It is not known why this is so and the specificity of Hib for the skin and subcutaneous tissue of the epiglottis remains an extraordinary trademark of Hib, particularly, in light of its specificity for humans.

1.5.2 Pathogenesis of invasive Hib disease

Most children exposed to Hib will experience no overt disease at all, since the majority will only harbor the organism as a harmless commensal in the throat (Michaels *et al.*, 1976). Disease, when it occurs, appears to occur within a short period (less than three days) after the initial encounter (Ryan and Falkow, 1994). It is not known why some individuals and not others develop Hib disease and the precise mechanism by which Hib gives rise to invasive disease is not fully understood. Nonetheless, many significant and

exciting developments in the past 25 - 30 years have increased our understanding of the pathogenesis of Hib disease and *H. influenzae* disease in general.

The accepted pathogenesis of invasive Hib disease can be said to involve three basic sequential stages: 1) colonisation of the nasopharynx, 2) bacteremia and survival in the blood stream, and 3) central nervous system invasion (although direct invasion of the meninges via contiguous spread from the nasopharynx is not precluded, clear evidence has been presented to show the meninges are invaded by a hematogenous route (Moxon and Murphy, 1978)). Pathogenesis begins with the pharynx as the portal of entry once penetration of the mucosal barrier has occurred. There is some evidence to suggest that there is a complex regulatory cascade coordinating capsular biosynthesis and adherence factors that act cooperatively in establishing the microbe within susceptible hosts (Ryan and Falkow, 1994).

Bacteremia with Hib is thought to result from direct invasion of the sub-mucosal blood vessels, rather than to develop via the regional lymph nodes (Rubin and Moxon, 1983). The probability of developing meningitis has been shown to be directly related to the magnitude of bacteremia (Moxon and Ostrow, 1977) and other types of Hib disease, such as epiglottitis, are thought to represent metastatic disease evolved from bacteremia (Kristensen, 1999). The organism has been observed to occasionally exhibit highly invasive behaviour in an *in vitro* adhesion invasion assay and it was proposed that sporadic invasivness may correlate with the unknown events that precede Hib bacteremia (Johnson *et al.*, 1996). To reach the bloodstream, this non-motile bacterium must pass through the basement membrane and sub-epithelial tissue, and enter the endothelium of a blood vessel. The manner by which this occurs remains under investigation.

The infant rat model of *H. influenzae* meningitis has been used to study the early pathogenic events of Hib disease and has played a substantial role in the identification of Hib virulence factors. Infant rats less than three weeks of age contracted meningitis after intranasal challenge with Hib (Moxon *et al.*, 1974) and an age dependent susceptibility to meningeal invasion and bacteremia was found (Moxon and Ostrow, 1977). The incidence of bacterial meningitis was directly related to the intensity of bacteremia regardless of the age of the rat (Moxon and Ostrow, 1977).

Evidence obtained in both animals and humans suggests that a preceding or concurrent respiratory infection may facilitate invasion of the bacterium. In experimental studies influenza A virus has been shown to increase both the frequency and the magnitude of Hib bacteremia on challenge with Hib in infant rats and the number of Hib organisms needed to produce invasive disease in rats were 100 fold reduced when infection with influenzae A virus was established (Kristensen, 1999). Moreover, family members of children with Hib disease reported symptoms of respiratory infection in the preceding four week period before onset of Hib infection significantly more often than controls (Kristensen, 1999).

A microbial pathogen is now recognised as a highly adapted microorganism that may cause disease (overt damage to a host) because its survival strategy includes a requirement for infection (persistence, usually by multiplication on or within another living organism) (Falkow, 1990). Disease can thus be seen to be an inadvertent outcome of a microbe's strategy for survival in the host (Falkow, 1994). Untreated Hib meningitis and epiglottitis are fatal in most cases. Since death of the host terminates transmission and propagation of organisms and, thus, is disadvantageous to the infecting bacterium, vascular invasion by Hib may be circumstantial or accidental rather the result of evolutionary advantage (Barbour, 1996).

1.5.3 Virulence factors of Hib

In the past two decades a revolution has occurred in our understanding of the basic pathogenic mechanisms of microorganisms including many developments made in the understanding of the molecular genetics of the pathogenesis of *H. influenzae*. It is now clear that mechanisms used by pathogenic bacteria that cause infection and disease are numerous and diverse and usually include a co-regulated group of complementary genetic properties that are suited for the interaction of a particular microorganism with a particular host (Falkow, 1990). The complex mechanisms that cause disease that are used by both encapsulated and non-encapsulated *H. influenzae* strains have now been broadly studied and valuable insight has been gained on the road of discovery towards understanding why Hib causes disease. Recently, Marrs *et al.* (2001) presented an excellent review of the current understanding of the molecular genetics of the factors that may be involved in the virulence of *H. influenzae* (Marrs *et al.*, 2001). This section (1.5.3) includes a brief description of the factors that are currently considered to be

among the putative virulence factors of Hib. These include the polysaccharide capsule, adhesion factors, lipooligosaccharide, heme uptake, haemocin, peptidoglycan, IgA proteases, *H. influenzae* genetic island 1, tryptophanase, antibiotic resistance and metabolic and physiologic factors. A list of the putative virulence factors with references is found in Table 1.5.

1.5.3.1 Capsule

The polysaccharide capsule and, particularly, the type b capsule has traditionally been considered the most important virulence factor of *H. influenzae* (Turk, 1984). Animal studies of isogenic strains transformed with capsule associated DNA confirmed earlier observations that type b strains were more virulent than type a strains, which were more virulent than other capsular types (Zwahlen *et al.*, 1989).

The capsular polysaccharide of Hib is a linear polymer of ribose, ribitol, and phosphate (Crisel *et al.*, 1975) called polyribosylribitol phosphate (PRP). The host's immune system is targeted against PRP and considerable evidence indicates that antibody to PRP is a principal protective host factor (Granoff and Munson, 1986). Consequently, this is the basis of an effective vaccine.

The bacterial capsule aids colonisation as it mediates resistance to complement and protects against phagocytosis which is advantageous for its invasivness (Moxon and Vaughn, 1981; Sutton *et al.*, 1982; Zwahlen *et al.*, 1989; Moxon and Kroll, 1990; Noel *et al.*, 1992). Interestingly, it has also been shown to impede interactions with epithelial cells *in vitro* (St. Geme and Falkow, 1991).

All encapsulated strains possess the *cap* genetic locus, a segment of ~17 kb comprised of a region that shows serotype specificity flanked by regions sharing homology among all capsular types. In Hib, with very few exceptions, the *cap* locus contains an unstable direct repeat of the ~17 kb segment flanking an ~1 kb bridge region containing the gene *bex*A (Kroll, 1992). The tandem duplication of the *cap* locus in type b strains is characterised by a mutation in the *bex*A gene of one copy of the duplication. This results in an unstable gene region, with relatively high potential for recombination events. During recombination of the *cap* locus, mutants may be formed that possess a single copy of the *cap* locus containing the *bex*A mutation that results in non-encapsulated mutants. Thus, reduction to a single copy state renders the organism capsule deficient

and poorly able to survive *in vivo* compared to capsule proficient organisms (Preston and Apicella, 1999). On the other hand, such a recombination event may generate mutants possessing additional (n) copies of the *cap* locus with n-1 copies of *bex*A. Such mutants are hyper-encapsulated and show increased pathogenicity in the infant rat model (Kroll and Moxon, 1988).

A significant number (~35%) of invasive type b clinical isolates contain amplified capsule gene sequences possessing three, four, or even five copies of the Cap b repeat (Hoiseth *et al.*, 1992; Corn *et al.*, 1993). The biologic significance of amplified capsule genes remains speculative but as the capsule probably facilitates the survival of the organism, it is plausible that an increase in production of capsular polysaccharide would enhance the virulence of Hib. Extra copies of the repeat can be lost on a single subculture suggesting the amplified state is a disadvantage for growth *in vitro*; lacking positive selection for increased capsule production, the biosynthesis of excess PRP may be unfavourable because it places an excessive energy burden on the organism (Corn *et al.*, 1993). The fact that most clinical isolates do not appear to be amplified implies that two copies of the Cap b repeat, the tandem duplication, are sufficient to cause invasive disease.

1.5.3.2 Adhesion factors

Successful pharyngeal colonisation, the first vital step in the disease process of Hib after exposure, requires that an organism overcome the mucociliary escalator of the respiratory tract. One strategy employed by *H. influenzae* to surmount mucociliary clearance involves attachment to the respiratory epithelium. Both pilus and non-pilus attachment factors appear to facilitate this. In 1982, Guerina, *et al.* (1982) and Pichichero, *et al.* (1982) independently reported the presence of long peritrichous pili in isolates of Hib (Guerina *et al.*, 1982; Pichichero *et al.*, 1982). Both groups noted piliation was correlated with an increased attachment to oropharyngeal epithelial cells and was also associated with a capacity to agglutinate human erythrocytes. It has been shown that pili expression is regulated by reversible phase variation believed to be roughly 10^{-3} to 10^{-4} in both directions (Farley *et al.*, 1990). Generally, strains isolated from invasive disease do not express pili, while those colonising the nasopharynx do (Marrs *et al.*, 2001).

At least 14 serological groups of hemagglutinating pili have been distinguished among both Hib and non-capsulate isolates (St. Geme, 1994). While the majority of Hib isolates express pili belonging to one serotype (serotype 4), non-capsulate isolates exhibit considerable diversity in this respect (St. Geme, 1994). In one study, sera prepared against purified pili from 2 strains of Hib did not recognise any of 12 piliated non-capsulate strains (Gilsdorf *et al.*, 1992). Most strains of Hib contain a pilus gene cluster composed of five genes, *hif*ABCDE (van Ham *et al.*, 1994; Marrs *et al.*, 2001).

Several studies suggest the presence of adhesins distinct from hemagglutinating pili among type b isolates but the identity of a second adhesin of Hib remains obscure while at least two adhesive high-molecular-weight proteins (HMW1, HMW2) expressed by non-capsulate isolates have been described (St. Geme, 1994). *H. influenzae* pili have been recently reviewed (Gilsdorf *et al.*, 1997).

Other factors shown to be involved with Hib adherence include fibrils, outer membrane proteins P2, P5 and P5-related fimbrin, and lipooligosaccharide (Marrs *et al.*, 2001). Additionally, fibrils and protein species designated Hia, Hap, and PE-binding adhesin have been found in NCHi (Marrs *et al.*, 2001).

1.5.3.3 Lipooligosaccharide (endotoxin[1])

Lipooligosaccharide (LOS), an essential outer membrane component, has been shown by several independent investigations to be a major contributory factor to the pathogenicity of *H. influenzae* (Kimura and Hansen, 1986; Zwahlen *et al.*, 1986; Maskell *et al.*, 1992). Structurally, LOS has a core of oligosaccharide but lacks the O-antigen side chains characteristic of the enterobacterial lipopolysaccharide. Oligosaccharides are exposed on the surface and are antigenic and they differ both within and between strains. Other characteristics of *H. influenzae* LOS include its endotoxic effects (Flesher and Insel, 1978), strain (Inzana, 1983) and phase variability (Kimura and Hansen, 1986; Weiser *et al.*, 1990) and an ability to stimulate an inflammatory reaction (Marrs *et al.*, 2001).

---

[1] ***Historical Note*** Rickard Pfeiffer, the discoverer of *H. influenzae*, was one of two scientists (Eugenio Centanni was the other) who independently studied the same heat stable toxin in the late 1800s. Pfeiffer noticed that *Vibrio cholerae* synthesised a heat resistant substance that was not actively extruded by the bacterium. He added the Greek word for 'within' to toxin and named his discovery endotoxin. While the term is still in use, it was shown that endotoxin resides on the surface of bacteria, not in the interior. (Rietschel and Brade, 1992)

A wide variety of glyco-modifications to the core LOS molecule have been detected and individual strains can also dynamically alter the expression of these terminal consitutents by phase variation of genes (*lic1A*, *lic2A*, *lic3A*, *IgtC*) within chromosomal loci that encode LOS biosynthetic pathway enzymes; this is presumably an adaptive strategy for subverting the immune system (Marrs *et al.*, 2001). Molecular mimicry of LOS to common eukaryotic glycolipid structures confers an ability to adhere to epithelial cells and as such would represent an important virulence determinant (High *et al.*, 1993). It has been shown that both serotypable and non-serotypable strains have the capacity to express LOS glycoforms with terminal sialic moieties that confer a relative resistance to the bactericidal activity of normal human serum (Marrs *et al.*, 2001).

Studies to identify *H. influenzae* modulins or immunomodulators have focused on LOS. Though endotoxemia is not a prominent feature of Hib disease (Ryan and Falkow, 1994), there is *in vivo* evidence for the involvement of LOS in the activation of the immune system and the development of inflammation (Marrs *et al.*, 2001; Miyazaki *et al.*, 2001). Its role in generating cerebrospinal fluid leukocytosis, protein influx, and blood-brain barrier permeability is well-characterised (Burroughs *et al.*, 1993).

1.5.3.4 Heme uptake

H. influenzae is nearly unique among human pathogens in its absolute requirement for exogenously supplied heme (a term used here to describe all iron-protoporphyrin complexes). The ability to circumvent the iron-withholding tactics of the human host by competing effectively for heme *in vivo*, is, presumably, an important determinant of the virulence of Hib. Initial studies on heme acquisition systems identified several cell envelope components that may be involved in its uptake while providing evidence that multiple iron uptake systems are used by Hib (Hanson *et al.*, 1992; Hanson *et al.*, 1992; Lee, 1992). Molecular genetic studies have shown that at least two cell surface proteins, Tbp1 and Tbp2, encoded by the genes *tbp*A and *tbp*B, respectively, are needed for transferrin binding capacity; three genes (*hit*A/*fbp*A, *hit*B/*fbp*B, *hit*C/*fbp*C) are necessary for the utilisation of iron bound to transferrin or iron chelates; heme bound to hemopexin is utilised via the *hux*ABC operon; and, four proteins (HgpA, HgpB, HgpC, HhuA) have been described that bind either hemoglobin or hemoglobin-haptoglobin complexes (Marrs *et al.*, 2001). A functional TonB protein has been shown to be a requirement for both heme utilisation and virulence (Jarosik *et al.*, 1994; Jarosik *et al.*, 1995).

1.5.3.5 Haemocin

Over 90% of Hib strains produce the bacteriocin haemocin. Haemocin is toxic to virtually all non-type b strains of *H. influenzae* (LiPuma *et al.*, 1990). The notable absence of haemocin production by NCHi or other encapsulated types suggests that it may play a role in the pathogenesis of infections (Murley *et al.*, 1998). It is thought that haemocin may contribute to the ability of Hib strains to effectively compete with NCHi strains in nasopharyngeal colonisation, an important early event in the pathogenesis of invasive Hib disease (LiPuma *et al.*, 1990). The gene encoding haemocin immunity, *hmcI*, and several genes in the *hmc* locus that have been cloned and sequenced are specific to type b strains (Murley *et al.*, 1997; Murley *et al.*, 1998).

1.5.3.6 Peptidoglycan

Despite sterilisation of the cerebrospinal fluid permanent neurologic sequelae, such as, deafness and seizures occur in many children surviving Hib meningitis. Thus, bacterial killing is not sufficient to assure a favorable outcome. Bacterial components released during antibiotic-induced killing transiently enhance the inflammatory response in the subarachnoid space, and down modulation of this inflammatory burst improves outcome (Burroughs *et al.*, 1993). Peptidoglycan, a major cell wall component, has been studied in regard to its contribution to this pathology of Hib meningitis. Hib peptidoglycan was shown to cause profound increases in blood-brain barrier permeability and cerebrospinal fluid white blood cell counts in an adult rat model (Roord *et al.*, 1994). Significant brain edema and protein influx with little leukocytosis was shown in a rabbit model of meningitis (Burroughs *et al.*, 1993; Roord *et al.*, 1994).

1.5.3.7 IgA proteases

The production of IgA proteases (Kilian *et al.*, 1979) is thought to facilitate invasion through the mucosal barrier (Plaut, 1983). Both non-typeable and serotypeable *H. influenzae* express type 1 and type 2 IgA1 proteases and cleavage products can be detected in nasopharyngeal secretions (Marrs *et al.*, 2001). Vitovski, *et al.* (2002) have very recently shown that IgA1 protease activity is higher in disease isolates of NCHi than in carriage isolates and their results provide quantitative evidence that IgA1 protease activity is an important virulence factor in NCHi mediated pathogenicity (Vitovski *et al.*, 2002). However, the role of IgA proteases in Hib virulence has not been established (Marrs *et al.*, 2001).

1.5.3.8 HiGI1

A 16 kb genetic locus, designated *Haemophilus influenzae* genetic island 1 (HiGI1), has been found to be conserved in Hib strains and absent from the majority of other strains studied (Chang *et al.*, 2000). That it is flanked by direct repeats, has regions different in G-C content from the rest of the genome, and evidence suggests that it may have resulted from phage-mediated transfer raise the potential that HiGI1 may be a virulence associated region (Chang *et al.*, 2000). Studies are under way to test the hypothesis that HiGI1 is a virulence determinant.

1.5.3.9 Tryptophanase

Among strains of *Haemophilus influenzae*, the ability to catabolise tryptophan has long been correlated with pathogenicity (Kilian, 1976). To begin to address the question as to whether this identifies tryptophanase as a virulence factor, the putative tryptophanase operon of Hib (strain Eagan) has been sequenced (Martin *et al.*, 1998). The Hib *tna* genes are not in the *H. influenzae* Rd KW20 genome, are flanked by direct repeats of an uptake signal sequence and appear to have been inserted as a mobile unit resembling a pathogenicity island (Martin *et al.*, 1998). Studies using gene disruption are currently being used to investigate if the *tna*A gene in the *tna* operon is a virulence determinate.

1.5.3.10 Antibiotic resistance

The emergence of multi-resistant phenotypes is a crisis in the medical management of Hib infections and contributes to its virulence. During the last two decades problems with antimicrobial resistance in invasive Hib has increased substantially and worldwide, though, it varies considerably between areas and is ameliorated in those areas where widespread vaccination has caused a significant decrease in the incidence of Hib (Kristensen, 1999).

1.5.3.11 Metabolic and physiologic factors

In contrast to the study of classic virulence factors, an effort has also been made to determine which metabolic and physiologic bacterial factors are essential for the survival and growth of *H. influenzae in vivo*. In particular, a genome-scale phenotypic analysis made possible by the availability of the complete genome sequence of *H. influenzae* Rd KW20 led to an estimate that 38% of *H. influenzae* genes are critical for growth or viability of colonies on rich medium (Akerley *et al.*, 2002). A comprehensive

Table 1. 5 Putative virulence factors of *H. influenzae* type b

Factor	Virulence Associated Action(s)	References
Polysaccharide capsule	Mediates adherence to host cells; inhibits phagocytosis and complement mediated lysis	Moxon and Vaughn, 1981 Sutton *et al.*, 1982 Zwahlen *et al.*, 1989 Moxon and Kroll, 1990 Noel, *et al.*, 1992
Lipooligosaccharide (the term lipopolysaccharide was used originally)	Mediates adherence to host cells; diverse types mimic common eukaryotic glycolipids and inhibit host immune defenses; antigenic phase variable expression inhibits host immune defenses; has toxic effect on ciliated respiratory cells; induces cerebrospinal fluid leukocytosis, protein influx and blood-brain permeability	Flesher and Insel, 1978 Inzana, 1983 Kimura and Hansen, 1986 Maskell *et al.*, 1990 Weiser, *et al.*, 1992 Miyazaki *et al.*, 2001 Marrs, *et al.*, 2001
Hemagglutinating pili /Fimbriae	Mediate adherence to host cells, however, subsequent reversion to a non-piliated state may promote invasivness	Farley, *et al.*, 1990 Gilsdorf, *et al.*, 1992 van Ham, *et al.*, 1994
IgA1 proteases	Inactivates the predominant secretory IgA species, IgA1 (Note: the role of IgA proteases in Hib virulence has not been established)	Kilian, *et al.*, 1979 Plaut, 1983 Marrs, *et al.*, 2001
Porin including OMP P2	Contributes to the signalling of the inflammatory cascade in the rat brain; mediates adherence to host cells	Galdiero, *et al.*, 2001
OMP P5	Mediates adherence to host cells; diversity contributes to antigenic drift; targetting of the carcinoembryonic antigen family of cell adhesion molecules may be of primary significance in colonization	Hill *et al.*, 2001
Peptidoglycan	Promotes inflammatory reaction; its most potent effect is a profound increase in blood-brain permeability and the induction of brain edema	Cundell, *et al.*, 1993 Burroughs, *et al.*, 1993 Roord, *et al.*, 1994
TonB protein	Promotes heme uptake and virulence expression	Jarosik, *et al.*, 1994
Ciliotoxin	Toxic effect on host	Ryan and Falkow, 1994
Haemocin	Toxic effect on other bacteria, including, most all NTHi and non-type b encapsulated bacteria	LiPuma, *et al.*, 1990 Murley, *et al.*, 1998
Haemophilus influenzae genetic island 1 (HiGI1)	HiGI1 is conserved in Hib strains; studies are underway to test the hypothesis that HiG1 is a virulence determinant	Chang, *et al.*, 2000
Tryptophanase	Ability to catabolize tryptophan is correlated with pathogenicity; studies are underway to test the hypothesis that the *H. influenzae tna*A operon is a virulence determinant	Martin, *et al.*, 1998
Antibiotic resistance	The emergence of multi-resistant phenotypes has occurred worldwide	Kristensen, 1999

list of factors that are essential *in vitro* under standard growth conditions have been identified from this work and these results provide information that be used to identify genes that are also essential during infection (Akerley *et al.*, 2002). An *in vitro* transposition system has also been described for identifying essential genes *of H. influenzae* (Reich *et al.*, 1999).

1.6 Epidemiology

In recent years there has been much progress in our understanding of the epidemiology of disease caused by Hib in developed countries. The quality of the data on the incidence of Hib infections in low- and middle-income countries varies but a broader picture is emerging of worldwide geographic and population differences in the incidence of Hib disease.

Before widespread immunisation against Hib the epidemiology of invasive Hib disease showed considerable variation through out the world (Table 1.6). In particular, rates of Hib disease 10-20 times that found in Western Europe (Takala *et al.*, 1989; Anderson *et al.*, 1995) and the United States (Murphy *et al.*, 1987; Shapiro and Ward, 1991) occurred in certain high risk populations, such as, Alaskan Eskimos (Ward *et al.*, 1981), native American Indians (Losonsky *et al.*, 1984) and Australian Aboriginals (Hanna, 1990). The estimates of disease incidence that are available from developing countries show that the incidence of invasive Hib disease can be similar to or more or less than that reported in industrialised areas but in general the incidence of Hib is greater in developing countries than in industrialised countries.

A remarkable finding is the huge differences reported with respect to age distribution of disease. At least 85% of all Hib infections occur in children less than 5 years of age and more than 65% occur in children less than 2 years of age. In non-industrialised areas and in native populations, children with Hib disease tend to be younger than in Western Europe, the United States, or Australian urban populations with a large proportion of the infants with Hib disease in these areas being younger than six months.

Major differences in mortality have also been recorded. Case fatality rates from Hib meningitis of 26-43%, 37%, and 57%, respectively, have been reported from Nigeria,

Gambia and Egypt while in Western Europe and the United States and Australia the mortality rate was around 2% (Campagne *et al.*, 1999; Kristensen, 1999).

The two most common manifestations of Hib disease are meningitis and epiglottitis in children. Though differing significantly in clinical presentation both diseases are life threatening and may cause harmful sequelae in survivors. In Western communities and indigenous populations it is striking that these infections tend to occur at different ages that vary according to the risk of Hib disease in the population. In populations where the incidence of Hib disease is high, epiglottitis is rare and meningitis predominates and occurs most often before 6 months of age. In contrast, where the incidence of Hib disease is lower epiglottitis is more common and usually occurs in children over 2 years of age and meningitis tends to occur predominantly in children over the age of 12 months. In developing countries where the incidence of Hib disease is high and most disease occurs before 6 months of age, pneumonia is commoner than meningitis and epiglottitis is not seen (Mulholland *et al.*, 1997).

1.6.1 Epidemiology of carriage

Carriage rates in different populations vary. In developed countries *H. influenzae* has been found to be carried by up to 80% of the population studied (Kuklinska and Kilian, 1984). Only a small percentage of the carriage strains are encapsulated and of these 2-4% are type b (Moxon, 1986). Most surveys in industrialised countries agree that nasopharyngeal or throat cultures recover Hib in 3% to 5% of young children with age being a prominent determinant of Hib carriage rates (Barbour, 1996). Higher rates up to 15% have been found in American Indian and Alaskan native populations and in some developing countries, such as, the Dominican Republic and the Gambia (Gomez *et al.*, 1998). On the other hand, the prevalence of Hib colonisation in children in Beijing, China has been reported to be 2% (Levine *et al.*, 2000). In Hong Kong, where there is an apparently low incidence of Hib disease, the carriage rate of Hib was zero in 621 Chinese children and 1.3% in 300 Vietnamese refugees (Lau *et al.*, 1998).

Table 1. 6 A selection of incidence and mortality data of invasive Hib disease from different worldwide areas

Country, Area or Population	Incidence /100,000 Children < 5 years old/year			Mortality (%)	Reference
	Meningitis	Epiglottitis	All		
Australia: Sydney	20	13	40	2.0	(McIntyre *et al.*, 1991)
Australia: Central Australia Aborigines, NT	273	0	991	8.3	(Hanna, 1990)
England and Wales	not stated	not stated	29	2.4	(Anderson *et al.*, 1995)
Finland	26	13	52	1.8	(Takala *et al.*, 1989)
USA: Minnesota	45	not stated	67	2.8[a]	(Murphy *et al.*, 1987)
USA: Dallas County, TX	67	not stated	109	2.8[a]	(Murphy *et al.*, 1987)
USA: Apaches	254	not stated	not stated	not stated	(Losonsky *et al.*, 1984)
USA: Alaskan Eskimos	409	0	491	6.0	(Ward *et al.*, 1981)
USA: Alaskan Natives (Indians and Eskimos)	282	8	601	3.1[b]	(Ward *et al.*, 1986)
USA: Alaskan non-Natives	69	16	129	3.1[b]	(Ward *et al.*, 1986)
France: Central and Southwest	15	2	21	2.8	(Reinert *et al.*, 1993)
Ireland	not stated	not stated	25.4	not stated	(Fogarty *et al.*, 1995)
Israel	18	not stated	34	2.1	(Dagan, 1993)
Greece	not stated	not stated	12	not stated	(Tsolia *et al.*, 1998)
Spain: Barcelona	not stated	not stated	15.4	not stated	(Beni *et al.*, 1999)
Spain: Catalonia	not stated	not stated	6.4	not stated	(Dominguez *et al.*, 1999)
Brazil: Campinas region	not stated	not stated	17	not stated	(Weiss *et al.*, 2001)
Hong Kong	not stated	not stated	2.7	not stated	(Lau *et al.*, 1995)
Singapore	not stated	not stated	3.3	not stated	(Lee *et al.*, 2000)
Korea	not stated	not stated	similar to USA	not stated	(Lee, 1998)
Japan	not stated	not stated	4.7	not stated	(Kamiya *et al.*, 1998)
China: Hefei City	not stated	not stated	10.4	not stated	(Yang *et al.*, 1998)
Nigeria: Niamey	not stated	not stated	179	43	(Campagne *et al.*, 1999)
West Gambia	60	not stated	45	37	(Bijlmer *et al.*, 1990)

[a]combined mortality for Dallas County and Minnesota
[b]combined mortality for Alaskan natives and non-natives

Under crowded circumstances or in a closed population carriage rates may be substantially higher than in open populations (Barbour, 1996). Prior to the introduction of the Hib vaccine, carriage of Hib in infants in an Aboriginal population in which the organism is highly endemic was found to be at least 48% by age 6 months (Smith-Vaughan *et al.*, 1998).

Since most patients with Hib disease have not had contact with a person who had invasive disease, and the organism has no known reservoir outside humans, asymptomatic carriers have been recognised as the major source of Hib infection (Barbour, 1996). Carriage of Hib lasting more than six months has been reported (Michaels *et al.*, 1976; Michaels and Norden, 1977).

Pharyngeal carriage of Hib is important, not only, in the transmission of Hib, but also, in the pathogenesis of Hib disease (Moxon *et al.*, 1974) and the development of immunity to the bacterium (Bradshaw *et al.*, 1971). The remarkable success of current vaccination programs against Hib has been due in part to the effect of conjugate Hib vaccines in decreasing carriage of Hib (Peltola *et al.*, 1992; van Alphen *et al.*, 1994; Barbour *et al.*, 1995; Muhlemann *et al.*, 1996).

There is some evidence that the population structure of nasopharyngeal Hib is different from that of invasive Hib (Leaves *et al.*, 1994). And, organisms isolated from the pharynx appear to lack certain virulence attributes found in organisms that have disseminated into the blood stream of patients with invasive disease (Weiser, 1993).

1.6.2 Epidemiology of Hib disease in Australia (Table 1.7)

Before the introduction of conjugated vaccines, Hib had been the most important cause of meningitis in young children in Australia and similar industrialised countries. Among Caucasian populations in major urban areas of Australia the reported incidence of Hib disease ranged from 25 to 63.5 cases per 100,000 children under the age of five. Few cases were attributed to other encapsulated types or non-encapsulated *H. influenzae*.

As with North American Indians and Eskimo children in the pre-vaccine era Aboriginal children had very high incidence rates of *H. influenzae* infection, young ages at disease occurrence, and a spectrum of disease characterised by the absence of epiglottitis and

predominance of meningitis and pneumonia. A striking difference between North American Indian children and Aboriginal children is that, whereas, virtually all cases in the former group of children are caused by type b strains, 15% of the infections in Aboriginal children in the Northern Territory are not caused by type b strains (Hanna, 1990). In this regard, the epidemiology of invasive *H. influenzae* infections in Aboriginal children is similar to that in children in some developing countries where Hib predominates in meningitis in young children but encapsulated types other than b and NCHi are frequently found as pathogens (Gratten *et al.*, 1994).

In Australia, epiglottitis comprises 30-40% of the cases of invasive Hib disease in non-Aboriginal children and meningitis comprises around 50% (Gilbert *et al.*, 1990; McIntyre *et al.*, 1991; McGregor *et al.*, 1992) but, in Aboriginal children epiglottitis is virtually unknown (Hanna *et al.*, 1992). The absence of epiglottitis among Aborigines corresponds to the striking characteristic of age-related variation in the incidence of Hib disease. The higher the overall incidence of Hib disease is, the earlier the median age of illness and the greater relative frequency of meningitis. Aboriginal populations have a very high incidence of Hib disease that occurs at a mean age of 6 months. Correspondingly it has been found that epiglottitis rarely if ever occurs in these populations (Clements and Gilbert, 1989). Another pattern of Hib disease in Australia before the introduction of Hib vaccines was a significant difference in the incidence of epiglottitis among urban populations. The incidence of epiglottitis in Victoria (Gilbert *et al.*, 1990) and the Australian Capital Territory (McGregor *et al.*, 1992) was twice that in Sydney (McIntyre *et al.*, 1991) and Western Australia (Hanna *et al.*, 1992).

The intriguing issues concerning the epidemiology of Hib disease have prompted numerous studies to examine host factors and the characteristics of bacterial strains associated with Hib meningitis and epiglottitis (Takala *et al.*, 1987; Johnson *et al.*, 1993; Johnson *et al.*, 1996). Despite such efforts no evidence has emerged that provides definitive answers to the questions raised. It has been reported that some Hib outer membrane protein (OMP) subtypes are more likely to cause meningitis than epiglottitis (Takala *et al.*, 1987) but this finding has not been substantiated by later studies (Takala *et al.*, 1989; Clements *et al.*, 1992). To date no finding has defined the characteristics of Hib, if any, that are associated with its unique epidemiology.

Table 1. 7 Incidence of Hib disease in Australia before the introduction of vaccines

Geographic Area	Years	Incidence /100,000 Children < 5 years old/year			Reference
		Meningitis	Epiglottitis	All	
NON-ABORIGINAL					
ACT	1984-1990	31	20	63.5	(McGregor *et al.*, 1992)
Victoria	1985-1987	25.4	22.7	58.5	(Gilbert *et al.*, 1990)
Sydney	1985-1987	20	1/3 of Hib disease	38.5	(McIntyre *et al.*, 1991)
Queensland	1990-1993	not stated	not stated	36.4	(Scott, 1993)
South Australia	1988-1991	25	not stated	25	(El Saadi and Cameron, 1993)
Western Australia	1984-1988	26.9	13.5	36.8	(Hanna and Wild, 1991)
Top End, NT	1985-1988	53	2 cases	59	(Hanna, 1990)
Central Australia, NT	1985-1988	215	2 cases	215	(Hanna, 1990)
ABORIGINAL					
Queensland	1990-1993	not stated	not stated	116.6	(Scott, 1993)
Western Australia	1984-1988	149.6	no cases	261.8	(Hanna and Wild, 1991)
Top End, NT	1985-1988	159	no cases	237	(Hanna, 1990)
Central Australia, NT	1984-1988	273	no cases	991	(Hanna, 1990)

1.7 Hib Vaccines

Most individuals acquire protective immunity in the first few years of life without ever having developed invasive Hib disease. This naturally acquired immunity probably develops as a result of pharyngeal colonisation with Hib, as well as, colonisation of the gut with normal enteric flora that are antigenically cross-reactive with Hib (Shapiro and Ward, 1991).

About seventy years ago it was recognised that inadequate levels of a 'protective bactericidal activity in the blood' in young children played a major role in the development of disease (Fothergill and Wright, 1933). It was later shown that

antibodies to the type b capsular polysaccharide mediated this immune response (Turk, 1981). We now know that transplacental anti-capsular antibodies in the newborn are lost within the first few months of life but reappear later in older children and adults, probably as a result of exposure to Hib or to other microbial antigens that engender cross-reactive antibodies. Loss of transplacental antibodies in the early months of life is coincident with the marked increase in the risk of Hib disease. The premise underlying the administration of vaccines during the first 6 months of life is to eliminate the window of age-related susceptibility between about 6 months (earlier in indigenous populations) and 2-4 years of age.

Although natural immunity to Hib undoubtedly involves immunity to several surface antigens of the organism, antibody to the type b polysaccharide, polyribosylribitol phosphate (PRP), appears to be of primary importance and the protection afforded by anti-capsular antibodies have been exploited in the development of Hib vaccine. PRP was purified for use as a vaccine immunogen in 1970 and vaccination against Hib was first attempted in a large field trial in Finland in 1974 (Peltola *et al.*, 1977). From this study several important conclusions were drawn: 1) adverse effects were minor, 2) antibodies were induced only in children above 18 months of age, 3) no booster effect was seen, 4) protective efficacy was 90% in children above 24 months and nil in children below 18 months, and 5) protection was related to the level of antibody against Hib capsular polysaccharide (Kristensen, 1999).

Like most polysaccharide vaccines, purified PRP is a relatively poor immunogen and does not provide immunologic memory in young children, hence, its inability to induce protective antibodies in children below 18 months. Subsequently, second generation vaccines have been developed in which PRP has been conjugated with various proteins to convert it to a thymus dependent antigen that is immunogenic in younger children and induces immunologic memory and an amnestic response to subsequent doses (Clements and Gilbert, 1989).

Four basic types of conjugate vaccines have been licensed in various countries. They are polysaccharide conjugated to diphtheria toxoid (PRP-D: 'ProHIBit', Connaught Laboratories), non-toxic oligosaccharide mutant diphtheria toxin (HbOC: 'HibTITER', Praxis-Lederle Biologics), outer membrane protein complex of *Neisseria meningitidis* (PRP-OMP or PRP-OMPC: 'PedvaxHIB', Merck, Sharp and Dohme), or tetanus toxoid

(PRP-T: Pasteur-Merieux). They differ, not only, in their carrier proteins, but also, in the structure and length of the polysaccharide molecule(s) bound to the protein, in the method of coupling the protein to the carbohydrate and in the ratio of protein to carbohydrate. All four are highly immunogenic in adults and older children but differ markedly in the immune responses they stimulate among infants and these differences appear to be predictive of efficacy, such that some vaccines might be more suitable in certain populations (Decker and Edwards, 1998). Decisions about conjugate Hib vaccines for the Australian schedule were based on local epidemiological data and extrapolation from studies in comparable populations elsewhere (Guthridge *et al.*, 2000). This led to the choice of PRP-OMPC for Aboriginal children and HbOC for non-Aboriginal children.

The conjugate Hib vaccines were introduced in Australia in 1992 and 1993 and a nationally funded program of infant vaccination was begun in July 1993. Since then, as in other areas where widespread immunisation has been introduced, the incidence of invasive disease has dropped dramatically and surveillance systems have noted a greater than 70% decline in invasive Hib disease in Australian children under the age of 5 years (Herceg, 1997; Bower *et al.*, 1998). In fact, where ever conjugate Hib vaccines have been introduced a dramatic reduction in Hib disease has been observed (Peltola *et al.*, 1992; Adams *et al.*, 1993; Booy *et al.*, 1994; van Alphen *et al.*, 1994). The impressive impact of Hib vaccines has led to the speculation that in areas where vaccination has been introduced, particularly where the vaccine coverage is high, it is now likely that Hib disease can be eliminated (Heath, 1998).

The process of eliminating Hib disease is helped by the eradication of Hib from the nasopharynx, a finding unique among bacterial infections (Peltola, 1993). That Hib conjugate vaccines protect against the acquisition of carriage has been established in epidemiologic terms, however, molecular knowledge about the mechanism by which conjugate vaccines reduce mucosal colonisation and the role of mucosal IgA antibodies in host defense against Hib is limited (Barbour, 1996). It is also not clear how large the impact of decreased carriage is among unvaccinated children for whom a reduced incidence of Hib disease has been observed (Madore, 1996).

While it is clear that Hib vaccination confers protection, as Hib disease incidence has plummeted where it is available, the picture emerging is not that simple. Reports of

increased carriage rates and the re-emergence of Hib after the introduction of vaccination and evidence of replacement disease contribute to the need to monitor vaccine efficacy, *H. influenzae* epidemiology and the distribution of genetic types of Hib. Continued high carriage rates averaging 9.3% but, as high as 22%, reported among Alaskan Natives in remote villages is postulated to have caused a resurgence in Hib disease in Alaska (Galil *et al.*, 1999). A sharp decline in incidence after the introduction of Hib conjugate vaccine PRP-OMP was observed in Alaskan Natives, however, after switching to a different Hib conjugate vaccine, DTP-HbOC, a re-emergence of Hib disease occurred suggesting continued transmission of Hib despite widespread vaccination (Galil *et al.*, 1999).

A different scenario has emerged in another Native American population. In a group who have historically had a very high rate of Hib disease, the Navajo Indian population in the USA, incidence rates have remained about 20 times higher than in the general USA population despite dropping tremendously (Scheifele, 2001). However, in a two year period only 1 Navajo child out of 380 was colonised with Hib and the locus and reservoir of colonisation remains unknown (Scheifele, 2001).

In Australia, prior to the introduction of vaccination, the carriage rates among Aboriginal infants in a remote community were at least 42.8% by age 6 months; these rates fell during the first 6 months of vaccine use and reached zero in 1994; but, subsequently, Hib carriage rates in infants in the same population have risen (Smith-Vaughan *et al.*, 1998). This is of some concern, particularly, in light of the re-emergence of Hib disease in Alaskan Natives. Relatively lower serologic responses to PRP-OMP vaccine in Aboriginal infants compared to Caucasian infants may contribute to the persistance of Hib colonisation (Guthridge *et al.*, 2000). The data also suggest that Aboriginal children may have a disproportionate proportion of reported deaths and vaccine failures (Herceg, 1997).

There was concern that as reduction of Hib disease progressed there would be replacement disease or replacement carriage with other types of capsulate or non-capsulate *Haemophilus* and that other serotypes might acquire additional virulence traits and emerge as important pathogens (Slack, 1995; Muhlemann *et al.*, 1996). It has also been speculated that the broad use of Hib vaccines may lead to an increase in invasive disease due to non-capsulated mutants of Hib since they are not susceptible to antibody

elicited by the vaccine (Kroll *et al.*, 1991). It is still not well known how the widespread use of vaccines influence the distribution of Hib subtypes.

Until recently evidence of replacement has not been observed (Barbour *et al.*, 1995) (Melville *et al.*, 1994; Madore, 1996), though, slight increases in NCHi disease have been reported (Muhlemann *et al.*, 1996). The predominant circulating Hib clones were shown to have disappeared during mass vaccination in Finland (van Alphen *et al.*, 1992). Then an unusual cluster of 5 cases of severe disease that resembles the clinical and epidemiologic features of Hib disease but was caused by *H. influenzae* type a occurred in a 10 month period in 1998-1999 in Utah (Adderson *et al.*, 2001). Three cases were attributable to a unique strain that possesses a deletion mutation associated with invasive strains of Hib but not normally found in non-invasive strains (Adderson *et al.*, 2001). The evidence of the possibility of the emergence of a new, virulent strain of *H. influenzae* with the virulence characteristics of Hib is of some concern.

Besides the issues noted here, it must be remembered that vaccine failures do occur (Scheifele *et al.*, 1996; Booy *et al.*, 1997; Breukels *et al.*, 2001) and that *H. influenzae* remains a common pathogen in adult patients in whom at least half of invasive infections are caused by Hib and other encapsulated types (Strausbaugh, 1997).

Investigation of carriage rates and the possible re-emergence of Hib disease will require long term surveillance of Hib disease. The need to monitor Hib isolates for the potential effects of the vaccine on clone selection or the replacement of Hib with other cap types is warranted. Such monitoring should include the genetic characterisation of carriage isolates, vaccine breakthrough strains and all strains from serious disease. To do this, techniques able to distinguish between Hib and the other serotypes and between capsule deficient type b and truly non-capsulate strains are essential.

1.8 Phenotypic and Genotypic Methods Used to Characterise Hib

Many questions about the epidemiology and population structure of Hib will be answered as we increase our understanding of the diversity that exists among Hib isolates. This section describes the variety of phenotypic and genotypic methods listed in Table 1.8 that have been used to characterise Hib and the general usefulness of the information derived from them for surveillance, epidemiological and population

structure studies. Discussion on the population structure and genetic diversity of Hib derived from studies using these methods is located in section 1.9.

Table 1. 8 Some Typing Systems Used to Characterise *H. influenzae* and Hib

Traditional Phenotypic Systems

Biochemical reactions (biotyping)
Capsular serotyping
β-lactamase production
Antibiotic sensitivity patterns

More Powerful Phenotypic Systems

Outer membrane protein typing
Lipooligosaccharide typing
Whole cell polypeptide typing
Multilocus enzyme electrophoresis typing

Molecular Typing Systems

Plasmid profiles
Polymerase chain reaction based methods:
- PCR-restriction fragment length polymorphism
- Randomly amplified polymorphic DNA
- Enterobacterial repetitive intergenic consensus sequences
- Long-PCR-ribotyping

DNA restriction fragment length polymorphism patterns derived from:
- Plasmid restriction endonuclease analysis
- Chromosomal restriction endonuclease analysis
- Pulsed field gel electrophoresis

Ribotyping
Southern hybridization
DNA sequencing of a particular locus
Whole genome sequence typing - *has begun*
Multilocus sequence typing - *future*

1.8.1 Classic phenotypic typing systems

Traditionally, methods used to discriminate among bacterial isolates in the epidemiological analysis of outbreaks of pathogenic bacteria have relied on what we now refer to as the traditional or classic phenotypic typing systems (Table 1.8). These systems detect characteristics expressed by microorganisms, such as, antibiotic and phage susceptibility patterns and biochemical and immunological reactions.

Because phenotypic typing systems detect characteristics of isolates that cannot be related to allelic variation at specific gene loci, it cannot be concluded that isolates sharing one or more phenotypic features are from the same or a closely related clone. Conversely, random mutations may also cloud interpretation of phenotypic data among variants derived from a single clone that exhibit different phenotypes due to a random mutation but that are otherwise genetically similar.

For these reasons using a single phenotypic characteristic as a critical determinant in any identification system is particularly hazardous. Usually several characteristics are

evaluated but this does not overcome the significant limitations in the use of phenotypic characteristics. They do not provide the information needed for the study of population structure because they are not able to detect the great diversity inherent in bacterial species and do not identify genotypes. Nonetheless, phenotypic methods are still very important in the primary identification of bacteria and can be useful in epidemiology studies.

1.8.1.1 Primary identification of Hib

The primary identification of Hib is based on the observation of typical colonial morphology on chocolate agar, gram stain morphology, and a characteristic mousy odor. The standard work-up for an isolate that exhibits microscopic and colonial properties suggestive of *Haemophilus* species is to determine the organism's X and V factor requirements (Table 1.1). Porphyrin-based testing methods (which detect strains of *Haemophilus* capable of synthesising heme, a property not shared by strains requiring exogenous X factor) represent the most reliable means for identifying *Haemophilus* species (Munson *et al.*, 2002). Key biochemical tests can be used to identify *Haemophilus* to the species level. Specifically, *H. influenzae* will ferment glucose but not sucrose, lactose, or mannose. It will not hemolyse horse blood agar and is catalase positive. Capsular serotyping techniques or capsular genotyping techniques are used to determine the capsular type.

1.8.1.2 Biotyping using Kilian's system

As previously mentioned in section 1.2, Kilian's system of biotyping based on indole production, and urease and ornithine decarboxylase activities is used to subdivide *H. influenzae* into eight biotypes (Table 1.2). The data that have accumulated from biotyping studies suggest that *H. influenzae* is biochemically heterogenous and initially it was thought that there might be a relationship between biotype and ability to cause invasive disease. However, this possibility is probably no more than a reflection of the unequal distribution of type b strains among the biotypes (Turk, 1984). Eighty-nine percent of Hib strains, the most frequent cause of the more serious *H. influenzae* diseases, are biotype I (Granoff *et al.*, 1982) whereas the majority of clinical NCHi fall into biotypes I, II, III, and IV (Barenkamp *et al.*, 1982; Butt *et al.*, 1990).

Biotyping may be useful in providing presumptive evidence for the epidemiology of *Haemophilus influenzae* disease as certain biotypes have been shown to be associated with different types of infections, sources of isolation, antigenic properties, and antimicrobial resistance patterns (Koneman *et al.*, 1988). The sensitivity of biotypes to strain variation, however, is limited since only eight biotype categories are defined and some biotypes are found infrequently. Consequently, biotyping of *Haemophilus influenzae* strains has relatively poor discriminatory power and is not useful in population genetic studies. Among Hib strains, most of which are biotype I, the discriminatory power is very poor and biotyping is not very useful for determining genetic diversity in epidemiological studies.

1.8.1.3 Capsular serotyping

Capsular serotyping is most useful in the primary identification of *H. influenzae* isolates. Slide agglutination utilising serotype specific antiserum has been the most widely used method to determine serotypes but this method has been shown to be prone to error (Murphy and Apicella, 1987). A common pitfall appears to be serotyping a non-capsulate isolate as a typeable one. This may be due to strains of NCHi autoagglutinating and the 'halo error' of high expectation that an isolate from blood or cerebrospinal fluid will be type b, especially when type b antiserum is the primary reagent used (Murphy and Apicella, 1987). False negatives due to technical error may also occur. Reports of infections due to non-typeable strains falling over 30% since the introduction of Hib vaccines in England suggest that some reported NCHi are really type b (Melville *et al.*, 1994). In addition, serotyping cannot detect *cap* b^- mutants that do not express capsular polysaccharide. They are mistakenly identified as non-typeable.

Serotyping has little discriminatory power because most invasive strains are type b. The use of genotypic capsular typing systems has contributed to the precise identification of capsular and non-capsular *H. influenzae*, as well as, to the elucidation of genetic variation among Hib strains (Kroll *et al.*, 1991; Falla *et al.*, 1994).

1.8.1.4 Antibiotic susceptibility and β-lactamase production

Antibiotic sensitivity patterns and β-lactamase production may be helpful in discriminating isolates in outbreak situations but tend to be of limited use in studies of a large number of epidemiologically unrelated isolates. Furthermore, antibiotic

susceptibility patterns can change rapidly as an organism may become resistant and, conversely, lose resistance by multiple genetic mechanisms. The identification of new or unusual antibiotic sensitivity patterns among isolates from multiple patients may be the first clue to an outbreak but offers little utility for epidemiology studies (Arbeit, 1995). Molecular characterisation of antimicrobial resistance determinants and plasmids has contributed to our understanding of the epidemiology of antimicrobic resistance and the genetic relatedness of multiply resistant *H. influenzae* (Mendelman *et al.*, 1985; Campos *et al.*, 1989; Levy *et al.*, 1993).

1.8.2 More powerful phenotypic typing systems

In the last 25 years there have been substantial developments in the ability to assess microbial diversity through the study of proteins. Though relying on the expression of genetic information, protein and other chemical profiling techniques have proven to be useful methods for discriminating bacterial isolates and have been developed as marker systems for epidemiology studies and to investigate genetic diversity. The most common protein analyses include outer membrane protein typing, whole cell polypeptide typing, and multilocus enzyme electrophoresis. Multilocus enzyme electrophoresis, in particular, has been used effectively for analysing population structure. Various types of lipid, carbohydrate and other chemical profiling have also proved useful in detecting diversity. The more powerful phenotypic systems used in the analysis of *Haemophilus influenzae* are described here.

1.8.2.1 Outer membrane protein typing

Useful phenotypic and epidemiologic information can be obtained from comparing outer membrane protein (OMP) profiles. The OMPs are isolated from the cell in insoluble fractions and the identification of principle proteins is established by determining their molecular weights on stained electrophoretic gels. A selective solubility of the cytoplasmic membrane proteins in a particular detergent is exploited in the isolation of OMPs. The conditionally insoluble membranes contain integral and likely some peripheral proteins which are released by relatively small changes in pH or ionic strength and are referred to as outer membrane proteins, and more appropriately, detergent insoluble membrane proteins (Carlone *et al.*, 1986). Sodium dodecyl sulphate-polyacrylamide gel electrophoresis of OMP preparations of Hib resolve a number of proteins that have been used to develop sub-classification schemes (Loeb and Smith,

1980; Barenkamp *et al.*, 1981; van Alphen *et al.*, 1983). Among both encapsulated and non-capsulate *H. influenzae* the outer membrane protein component contains up to 36 proteins of which 6 represent the major protein content (Loeb and Smith, 1980). The major proteins have molecular weights between 50,000 and 15,000 and are labelled P1 to P6 or a to f in order of decreasing molecular mass (Foxwell *et al.*, 1998). Overviews of the characteristics of major and minor proteins are found in Foxwell, *et al.* (1998) and Gilsdorf's reviews on *H. influenzae* (Foxwell *et al.*, 1998; Gilsdorf, 1998).

The usefulness of OMP subtyping varies. Initially, Granoff, *et al.* (1982) identified 21 distinctive OMP subtypes among Hib isolates from 256 patients in 22 states in the USA. They found that three subtypes accounted for nearly three-quarters of the isolates and six subtypes were responsible for 89% of infections (Granoff *et al.*, 1982). That a limited number of types accounted for most of the isolates suggested a small number of distinctive clones were responsible for Hib disease. In contrast, studies in the Netherlands showed that a single OMP subtype predominated in Western Europe (van Alphen *et al.*, 1983; van Alphen *et al.*, 1987). Similarly, a single OMP subtype was responsible for 83% of Hib infections in Victoria (Clements *et al.*, 1992) and 76% in Western Australia and the Northern Territory (Hansman and Lawrence, 1993). Thus, OMP subtyping in those regions is of limited value for epidemiological purposes.

While OMP typing indicates that there is considerable genetic diversity of Hib in some populations it does not measure the extent of variation in the genome and it has a limited ability to discriminate among Hib unless an infrequent subtype is involved. In order to answer questions about genetic relatedness a more powerful typing method is needed. Despite technical and practical limitations OMP subtyping is frequently used to distinguish Hib isolates and its widespread use has contributed to our understanding of the diversity of Hib.

1.8.2.2 Whole cell polypeptide typing

The technique of whole cell polypeptide typing characterises the molecular weights and mobility of the whole cell polypeptide contents in a stained electrophoretic gel using sodium dodecyl sulphate-polyacrylamide gel electrophoresis. It has been assessed to be more discriminatory than biotyping and outer membrane protein typing for examining the genetic diversity of Hib and NCHi because of the large number of bands found among the isolates studied; in two reports the average number of bands found per

isolate was 39 and 48 (Paterson *et al.*, 1987; Pennington and Freebairn, 1989). While the large number of markers, a feature necessary for the analysis of population structure, increases the discriminatory power of whole cell polypeptide typing, it makes it difficult to compare large numbers of isolates. Thus, whole cell polypeptide typing is probably best suited for small-scale epidemiological studies. The disadvantages of any polyacrylamide gel technique are inherent in the whole cell polypeptide analysis, the most notable being problems associated with inter-gel variability. The study of the genetic diversity of Hib and NCHi in North East Scotland resulting in the identification of clones and the estimation of the total number of clones existing in that geographical region with comparison to isolates from England, Sweden, the Netherlands, and the USA is the primary body of work that has been reported for *H. influenzae* using whole cell polypeptide typing (Paterson *et al.*, 1987; Pennington and Freebairn, 1989; Bruce and Jordens, 1991; Bruce and Pennington, 1991).

1.8.2.3 Lipooligosaccharide typing

Characterisation of Hib lipooligosaccharide (LOS) by sodium dodecyl sulphate-polyacrylamide gel electrophoresis revealed micro-heterogeneity among the LOS molecules of different Hib strains, which permitted the development of a limited system for subtyping Hib strains based on their LOS profile (Kimura and Hansen, 1986). The antigenic heterogeneity results primarily from differences in lipid A as well as differences in its core polysaccharide or short carbohydrate side chains (Gilsdorf, 1998). Originally 11 LOS types were described (Inzana, 1983). However, a paucity of heterogeneity among the cell surface exposed antigenic determinants of Hib LOS was confirmed in several studies including those by van Alphen, *et al.* (1983) who reported finding 4 different LOS serotypes among 80 strains of Hib in the Netherlands and Gulig, *et al.* (1984) who described four antigenic groups among 126 strains in the USA (Kimura and Hansen, 1986). Due to the limited LOS antigenic diversity, LOS profiles have been used most often in conjunction with other typing systems to characterise the genetic variation of Hib. NCHi strains exhibit a more heterogenous population with respect to LOS demonstrating no fewer than 14 groups based on LOS composition (Foxwell *et al.*, 1998).

1.8.2.4 Multilocus enzyme electrophoresis

Multilocus enzyme electrophoresis (MLEE) is a robust technique that has been used to estimate genetic diversity and structure in natural populations of a variety of species of bacteria by the analysis of up to 25 or more different metabolic enzymes per isolate (Selander *et al.*, 1986). Basically, the number of shared (or identical) enzyme variants between pairs of isolates is used to determine their relatedness. Water soluble cellular enzymes are extracted and electrophoresed on a non-denaturing starch gel or cellulose acetate sheets followed by staining of the enzymes. Each electrophoretic variant is assigned a numeric score and each unique combination of scores is called an electrophoretic type (ET).

Because the rate of migration of a protein during electrophoresis is determined by its amino acid sequence, mobility variants can be directly equated with alleles at the corresponding structural gene locus. A large proportion of amino acid substitutions can be detected by this method. But, because some substitutions do not affect electrophoretic mobility, electromorphs may be sequentially heterogenous and at the level of the nucleotide sequence of the gene itself, there is even greater heterogeneity, owing primarily to silent substitutions (Selander *et al.*, 1986). Thus, it must be understood that the alleles recognised may actually be groups of isoalleles. Post-translational modification is a potential source of error in the application of MLEE but there is no evidence that post-translational modification of enzymes occurs frequently enough in any organism to seriously bias the estimates of genetic variation derived from the electrophoresis of proteins (Selander *et al.*, 1986).

Criticisms of MLEE are that: 1) it assays the genotype indirectly, so that much variation at the nucleotide level may go undetected because nucleotide substitutions do not necessarily change the amino acid composition; 2) changes in amino acid composition do not necessarily change the electrophoretic mobility of the protein and, as a consequence, alleles that are considered to be the same protein alleles from different individuals may represent different gene alleles, and 3) selection may be acting on the polymorphisms, so that anonymous DNA markers may give a very different picture from allozyme markers, presumably because the former are neutral and the latter are under some sort of selection (Taylor *et al.*, 1999). In other words only a minority of mutations are detected (those that alter the electrophoretic mobility of the enzyme), and

enzymes with the same electrophoretic mobility can be encoded by very different gene sequences.

As MLEE can be used to detect genetic changes that accumulate slowly and are likely to be selectively neutral it is suitable for long-term and global epidemiology studies (Yakubu *et al.*, 1999). MLEE offers limited applicability for epidemiological analysis of clinical isolates and, for practical considerations alone, the cost, relative labor intensity and complexity of the method has restricted its application to primarily large scale population structure studies carried out in research laboratories. In fact, the context in which we now view bacterial population structure is based largely on studies conducted using MLEE. MLEE data have been used to estimate the levels of single locus and multilocus genotypic variation in populations as well as the extent of genetic exchange within a population (Selander *et al.*, 1987). A comprehensive analysis of the genetic structure of *H. influenzae* based on 2,209 isolates from six continents has been carried out using this method (Musser *et al.*, 1990).

1.8.3 Genotypic typing systems

The shortcomings of phenotypically based typing methods have led to the development of genotypic typing systems that mine the tremendous variation found in bacterial DNA. In the past twenty-plus years there has been a virtual explosion in techniques used to analyse nucleic acids. Phenotypic methods are now broadly supplemented by genotyping methods that allow a more refined and detailed differentiation of closely related microorganisms on the DNA level.

Molecular techniques, such as, polymerase chain reaction derived methods, restriction enzyme digestion methods, Southern blot analysis, and DNA sequencing methods have become routine procedures in the study of genetic diversity. Ever more innovative molecular techniques are continuously being added to the molecular typing palette often displacing first generation genotypic methods. Plasmid profile analysis, probably the first DNA based technique used in epidemiology studies, falls into the category of traditional typing methods that are still useful but have been largely supplanted by the newer genotypic methods of discrimination.

It should be noted that the explosion of molecular techniques has revolutionised the application of the differentiation of microorganisms. The techniques that involve direct DNA-based analyses of chromosomal or extra-chromosomal genetic elements have proven to be powerful tools for discriminating bacterial isolates. Their use has directly contributed to the growth of several major fields of microbiological research. Taxonomy, population genetics, the study of evolutionary dynamics and phylogenetic relationships, and microbial epidemiology all rely on the ability to discriminate among genotypes. Many areas of applied microbiological research also depend on the measurement of genetic diversity, including, for example, the investigation of outbreaks of infectious disease or measuring genetic diversity in relation to pathogenicity, drug resistance or biodegradation.

Substantial complexities in interpreting the results of genotypic methods and applying the data to epidemiologic studies remain, as does, the challenge of standardising methods and establishing a common nomenclature. Great strides have been made in each of these areas and it is certainly an exciting time in what can be called "the era of burgeoning bioinformatics".

The utility of molecular typing methods varies. Each method has its technical and nucleic acid target-dependent limitations that should be taken into account with the choice of method depending on the nature of the questions to be answered. It is commonly recommended that "polyphasic" systems incorporating both different molecular targets and additional analytical procedures be employed in a molecular identification system (Harmsen *et al.*, 2001).

In the following sections (1.8.3.1 – 1.8.3.6) a number of representative genotypic methods including a comprehensive overview on pulsed field gel electrophoresis are described. The usefulness of these methods in discerning the genetic diversity of Hib is discussed in section 1.9.

1.8.3.1 Restriction fragment length polymorphism analysis

Restriction fragment length polymorphism (RFLP) analysis is a powerful method for detecting sequence dissimilarities by analysis of size and number of fragments generated by endonuclease digestion of DNA. Methods using direct visualisation of stained fragments involve digestion of total DNA with a suitable restriction enzyme,

followed by electrophoresis in agarose or acrylamide gels or by PFGE. The DNA molecule is sampled for sequences that are cleaved by a particular restriction endonuclease; fragments of the same length and number are generated from genomes of identical or similar sequence, but fragments of different length and number are generated when the genomic sequences are different. The proportion of fragments that have identical sizes provides an estimate of the sequence similarity between DNA molecules. The patterns, provided they are not excessively complex, can then be subjected to numerical analysis, such as, Dice coefficients and dendrograms, to give statistical estimates of sequence similarity.

The reliability of numerical analyses and associated graphical representations based on RFLP data has been questioned because each restriction fragment does not represent a separate distinct characteristic; rather, each fragment is defined by two restriction sites, each comprising 4, 6, or 8 nucleotides (Arbeit, 1995). For PFGE profiles, which represent the entire circular chromosome, the creation or loss of a single restriction site affects two fragments. Moreover, fragments of the same molecular size cannot be assumed to represent the same chromosomal DNA and may differ because of the insertion or deletion of extrachromosomal DNA (e.g., bacteriophages) (Arbeit, 1995). This has led to questions about the relationship between isolates with similar restriction patterns, in particular, whether it is valid to relate restriction fragment length conservation with overall DNA sequence conservation in bacteria (Hall, 1994). Despite these theoretical concerns the use of RFLPs to determine genetic relatedness is well accepted and in practice has been shown to produce reliable data concerning relatedness (Dr Adrian Gibbs, personal communication). In a study of 29 epidemiologically distinct Hib isolates PFGE RFLPs reflected evolutionary divergence consistent with that identified with MLEE (Arbeit *et al.*, 1990). The issue requires additional investigation but currently numerical analysis of RFLPs is widely used to determine genetic relatedness. See interpretation of PFGE RFLPs in this section.

1.8.3.2 Chromosomal restriction endonuclease analysis

Chromosomal restriction endonuclease analysis (REA) involves comparison of the number and the size of fragments produced by digestion of DNA with a frequent cutting restriction endonuclease that cuts DNA at a constant position within a specific recognition site usually composed of 4-6 bp. A limitation of REA of whole

chromosomes is the generation of a large number of fragments. The separation by conventional gel electrophoresis leads to a smear-like appearance of the DNA on the gel and the resulting band patterns are too complex to be properly compared. To master this problem, the method is combined with nucleic acid probes (e.g, ribotyping). The use of pulsed field gel electrophoresis and rare cutting endonucleases overcomes this problem as the number of fragments generated can be limited to a manageable number due to the ability of PFGE systems to separate relatively large pieces of DNA.

The advantages of REA are that it is universally applicable (in principle all isolates may be typed by REA), sensitive because the whole genome is sampled, and relatively easy to perform. Its major disadvantage is the complexity of the band patterns. Forbes, et al. overcame this problem by dividing the fingerprints into sections on the basis of the size range of fragments and then comparing the fragments in a number of sections to establish similarities or differences (Forbes *et al.*, 1991).

1.8.3.3 Pulsed field gel electrophoresis

Conventional agarose gel electrophoresis is capable of separating DNA fragments with sizes up to 40-50 kilobases (kb) in length (Maniatis *et al.*, 1989). In the early 1980s Schwartz and co-workers developed an electrophoretic technique capable of resolving DNA molecules in excess of 2,000 kb (2.0 Mb) (Schwartz and Cantor, 1984) and in doing so ushered in pulsed field gel electrophoresis (PFGE), an advance in DNA technology of great significance; and, as the analysis of megabase sized DNA molecules has become routine the use of PFGE has become increasingly common.

PFGE utilises the ability of alternating electric fields to separate relatively large fragments of DNA up to megabases in size. The significance of this is that entire bacterial genomes can be digested using rare cutting endonucleases, which cut infrequently, and the resulting few large fragments can be separated using a PFGE system. So, theoretically, entire genomes can be compared by analysing restriction fragments. It has become a technique of central importance in the mapping and sequencing of entire genomes, prokaryotic or eukaryotic, as well as, a method of choice for discrimination of bacterial isolates in epidemiological studies. Because PFGE has been shown to be superior to most other methods for biochemical and molecular typing,

it is often considered to be the "gold standard" of molecular typing methods (Olive and Bean, 1999).

The term 'pulsed field gel electrophoresis' is a generic term referring to gel electrophoresis systems in which the electric field is switched, usually in a controlled, pulsatile fashion between two different directions with pulse times ranging from 0.1 to 1000 seconds or more. The variation is applied most commonly to field direction, but the pulsing may also be in field strength. By forcing DNA molecules to reorient periodically from one electric field direction to another PFGE overcomes the size limitations of conventional electrophoresis where all linear double stranded DNA molecules that are larger than about 40-50 kb migrate through agarose gels at the same rate. In PFGE systems, the larger the DNA molecule, the longer it takes to reorient. Shorter molecules can reorient more quickly and can therefore move through the gel faster than longer molecules. As a result mobility becomes a function of size and fragments ranging from 0.1 to 10 megabases can be efficiently separated.

A consistent and generally accepted theoretical basis for DNA behaviour during PFGE has yet to be described. Probably the most popular conceptual framework is described in Sambrook and Russell's third edition of Molecular Cloning: A Laboratory Manual (Sambrook and Russell, 2001). It proposes that the effect of alternating fields can be explained by considering that a DNA molecule stretches out or elongates and orients parallel to an electric field. It can then be moved through successive pores in the agarose gel by reptation (the movement of large molecules believed to assume an 'end-on' configuration to the gel matrix as they wriggle in a reptile-like mode through the gel interstitial spaces). When the field direction changes, the DNA molecule must reorient before it can move in the new direction. As the field direction regularly changes, the DNA must regularly change directions. The larger the molecule, the more time required to reorient and turn corners, leaving proportionately less time (within each pulse) available to move in the new direction.

A variety of different types of apparatus have been used successfully for PFGE. All are based on the approach to find the optimum electrode geometry which, coupled with appropriate pulses, would give reliable separation of large DNA molecules. Several of the more popular designs are described in Molecular Cloning: A Laboratory Manual (Sambrook and Russell, 2001).

Contour-clamped homogeneous-field (CHEF) electrophoresis is the most common type of PFGE now used. In CHEF PFGE the electrode configuration and the electrical switching are controlled in such a way as to generate electric fields that are homogeneous. The electrode arrays are hexagonal giving a field orientation, which is pulse-switched by 120°. The DNA molecules migrate with linear trajectories and fragment sizes can be compared accurately across multiple lanes in a gel so that size identity can be established. It also allows the accurate indexing of gel bands to transferred fragments on membrane filters in DNA hybridization analysis. And specific DNA molecules can be isolated by cutting gel bands from individual lanes.

There are five major practical determinants that must be considered in optimising the resolution of large DNA molecules on gels: *pulse time, field strength, alternating field angle, gel concentration (and agarose quality), and temperature* (Stewart *et al.*, 1994). In addition, special consideration must be given to *DNA preparation, the use of molecular weight markers and restriction enzymes.* These parameters are interdependent and require careful exploration to generate a successful PFGE. A great deal of this exploratory work has been done and is reported in accounts of PFGE but it is worth understanding the functional significance of variation of these parameters.

Pulse times should be greater for larger molecules. As a starting point from which to explore individual protocols for bacterial studies, pulse times of 1-2 seconds for <50 kb, 10 seconds for 50-200 kb, 60 seconds for 200-800 kb, 120 seconds for up to 1400 kb have been suggested (Stewart *et al.*, 1994). The use of ramped switching permits greater resolution of DNA fragments within a given size range. In ramping, the switch time at the beginning of the electrophoresis run is different from that at the end of the run. The resolution of a broad size range of fragments on a single gel can be optimised by programming consecutive blocks (each with a customised ramp of initial switch time, final switch time, and run time) for the electrophoresis run. Ramping and using a series of blocks of varying ramps is possible on most commercial CHEF systems.

Field strength or voltage, measured as V/cm, determines the rate of migration of charged particles through a gel. While increasing the field strength increases the size of the largest material resolved, there is a limit to the speed with which PFGE separations can be carried out because knotting and irreversible pinning of molecules to gel matrix fibers occurs at high voltages (Sambrook and Russell, 2001). Voltage effects are

particularly important for resolution of >2000 kb for which long, slow pulse times of 30-90 minutes at ~1-2 V/cm are required. For smaller fragments, typical of the fragments produced when bacterial genomes are digested with infrequent cutting enzymes, pulse times of 1-90 seconds at 6 V/cm are usually optimal.

Alternating field angle should be somewhere between 90° and 180° with 120° a common setting. In many systems this is an invariant parameter. It has been shown that decreasing the included angle from 120° to 94° increases the velocity of the DNA, with the mobilities of the large DNAs (>1 Mb) affected to a greater degree than the smaller DNAs (<1 Mb) (Bio-Rad, 1992).

Gel concentration and agarose quality are similar to those now used in conventional electrophoresis. Gel concentrations around 0.8-1.2% provide good band sharpness. The quality of agarose is important and the highest quality ultra-pure agaroses should be used. Agaroses with ultra-low electroendosmosis used with low ionic strength buffers can increase the mobilities of DNA molecules in PFGE by as much as 50% (Lai, 1991). It might be noted that it was the advent of PFGE with its requirement for contaminant free agarose and a reduction in its long run times that has led to the quality of agaroses that are now available for electrophoresis.

Temperature influences the rate at which DNA molecules relax and adopt new configurations when the electric field changes. DNA migrates faster at higher temperatures; however, higher temperatures also cause trapping at lower fragment sizes and significant band broadening (Sambrook and Russell, 2001). So, unlike conventional agarose electrophoresis, which is generally carried out at room temperature, PFGE is generally run at 4-15°C with temperatures of 12-14°C appearing to be optimal in increasing relaxation rates while at the same time avoiding band and lane distortion due to heat build-up (Stewart *et al.*, 1994).

DNA preparation for PFGE cannot be accomplished using standard procedures for DNA isolation in solution. Instead, to preserve the integrity of the large DNA molecules required for PFGE, DNA isolation is done by embedding live cells in agarose (Schwartz and Cantor, 1984). To prevent shear damage, procedures have been developed for extracting DNA from the cells embedded in small plugs (also called inserts) of low melting temperature agarose. The cells are lysed and other cellular macromolecules

(proteins, RNA, lipids) are degraded or solubilised by enzymatic digestion and detergent treatment directly in the agarose plugs. The 'cleaned' DNA embedded in the agarose plugs is subjected to buffer washes and digestion procedures for the restriction enzyme of choice and, subsequently, loaded into the loading wells of the gel. Prior quantitation of cells is an important criterion in plug production to ensure the amount of DNA on the gel is optimal because the mobility of DNA in PFGE is more sensitive to concentration than in conventional electrophoresis. Amounts must be restricted to <1-2 μg per lane (Sambrook and Russell, 2001).

Molecular weight markers of a high molecular weight are required for PFGE as size standards are critical for estimating sizes of fragments, particularly, in the comparison of RFLPs on different gels whether by visual inspection or with commercial software packages. The use of λ concatemers has become the standard for bacterial PFGE work. A series of markers evenly spaced over a wide range of molecular weights is obtained by the ligation of bacteriophage λ DNA into a nested series of concatemers. The principle behind the formation of the concatemers is as follows. The viral DNA is readily ligated by virtue of the 12 bp sticky ends of *cos* sites that occur at both ends of the bacteriophage genome. Lambda will connect to themselves at these *cos* sites and form multimers of its 48.5 kb monomer; thus the individual concatemers are multiples of this length. Digested bacterial isolates with known fragment sizes, such as, the *Sma*I digest of the chromosome of *S. aureus* ATCC 8325 used by Pattee to define the physical map of *S. aureus* are excellent size standards and preparation of the DNA serves as a useful quality control for the whole procedure (Maslow *et al.*, 1993).

Restriction endonucleases with infrequently occurring recognition sites are essential to exploit the resolving power of PFGE. Due to the non-random arrangement of base pairs in a genome, certain restriction sequences may be substantially under-represented and the base composition and/or sequenced DNA from a genome can be used to predict which restriction sequences will be rare. Often the selection of endonucleases is a matter of trial and error guided by the nucleotide composition (G+C content), size of the genome and the purpose of the PFGE. Two practical points to keep in mind when selecting restriction enzymes are: one, DNA of low G+C will cut infrequently when treated with an enzyme with a G+C rich recognition site and two, enzymes with 8 bp recognition sequences will cut less frequently than comparable 6 bp cutters. Restriction

enzymes that generate between 15-25 fragments with a broad size distribution are optimal.

Interpretation of PFGE restriction patterns. Analysing the restriction digest patterns resolved by PFGE to identify and differentiate strains is based on comparing the size and number of fragments between pairs of isolates. Tenover et al. have suggested and it is generally accepted that strains with identical PFGE patterns are considered to be the same strain and clonal. Strains with one or two band shifts consistent with a single genetic event (e.g., a point mutation resulting in the loss or gain of a restriction site, an insertion, a deletion, or a chromosomal inversion) are also considered to be closely related. If there are a very limited number of band shifts (4-6) which are not explained by a single genetic event, the strains are considered to be possibly related but distinct. Strains that differ at six or more bands are considered to represent independent strains, although strains with multiple or distinctive similarities may have a common ancestry (Maslow *et al.*, 1993; Tenover *et al.*, 1995).

Gel results can be photographed and the data can be stored by using one of several commercially available digital systems. Data analysis can be accomplished by using any of a number of commercially available software packages. With the aid of gel scanning and analysis software it is theoretically possible to create data banks of PFGE patterns for all organisms, enabling the creation of reference databases to which any new strain could be compared for identifying its phylogenetic relationship to similar strains (Olive and Bean, 1999). At this time inter-laboratory standardisation has not reached a level to allow the use of common type nomenclature or direct DNA pattern exchange to be used on a regular basis. However, surveillance networks are being been set up now using PFGE. These networks will serve as models for more extensive use of such standardisation. Several examples are noted here.

The most prominent PFGE network is PulseNet, an early warning system for outbreaks of foodborne diseases set up in 1996 by the Centers for Disease Control and Prevention (CDC). It consists of a national network of laboratories that performs PFGE on bacteria that may be foodborne. The network identifies and labels each PFGE pattern and permits rapid comparison of these patterns through an electronic database at CDC. The subsequent inclusion of PFGE analysis of other bacteria, such as *Neisseria meningitidis*, to PulseNet demonstrates its wider usefulness (Popovic *et al.*, 2001). Another network

is the Pneumococcal Molecular Epidemiology Network established in 1997 under the auspices of the International Union of Microbiological Societies with the aim of standardising the nomenclature and classification of resistant pneumococcal clones. This system identifies worldwide clones using PFGE with two other techniques (McGee *et al.*, 2001). A third example is the rapid standardised PFGE protocol for macrorestriction analysis of methicillin resistant *Staphylococcus aureus* (MRSA) that has been developed in Canada which allows for inter-center comparisons of MRSA fingerprints for national surveillance studies (Mulvey *et al.*, 2001)

Disadvantages of PFGE. Originally extensive use of PFGE was prevented by the long and tedious procedures required for isolating and cutting the genomic DNA by the agarose plug method, the requirement for large quantities of expensive enzymes and reagents, the complexity of experimentally optimising the electrophoretic conditions and the need for specialised equipment much more expensive than that required for PCR or southern hybridisation (Swaminathan and Matar, 1993). That is not the case now. New developments in instrumentation (CHEF DR™III, Bio-Rad) and simplification of the protocol address some of these problems (Matushek *et al.*, 1996; Gautom, 1997; Chang and Chui, 1998; Benson and Ferrieri, 2001).

Advantages of PFGE. PFGE is a highly reproducible and discriminating tool. It has been successfully applied to a broad range of different Gram-negative and Gram-positive organisms in epidemiological studies and outbreaks (Matushek *et al.*, 1996). These include *Escherichia coli*, *Neisseria meningitidis*, *Pseudomonas aeruginosa, Listeria monocytogenes*, *Bordetella pertussis*, *Legionella pneumophila*, *Staphylococcus aureus* and enterococcal isolates (Gautom, 1997). Its discriminatory power has been shown to be better than MLEE, ribotyping, and RAPD-PCR (refer to section 1.8.3.4) methods and is best used to identify the micro-variation that is required to distinguish between strains; it is suitable for typing both highly clonal and less clonal populations (Struelens, 1998)

Highly variable regions of the genome or individual loci can be identified and appropriate restriction enzymes chosen (Yakubu *et al.*, 1999). From a theoretical perspective PFGE is considered to be appropriate for detecting replication errors resulting from point mutation, recombination, and inversion and for detecting acquisition of extraneous elements through transformation, phage transduction or

infection and transposons (van Belkum *et al.*, 2001). The sensitivity of PFGE to detect genomic rearrangements makes appropriate interpretation of minor pattern differences a key to its correct application (Struelens, 1998).

Criticisms of PFGE. PFGE is now the "gold standard" for epidemiological typing and calculation of restriction pattern similarity coefficients and graphical display of pattern relatedness as dendrograms is also useful for interpretation, particularly for large scale studies (Struelens, 1998). Although this analysis has been criticised as invalid for phylogenetic inferences because DNA restriction fragment pattern variation is not due to independent events, it is supported by population analysis, e.g., of *Pseudomonas aeruginosa* (Musser *et al.*, 1995; Struelens, 1998). It is also a concern that fragments of the same size cannot be assumed to represent the same chromosomal region and, conversely, that restriction fragments of different sizes may represent the same chromosomal DNA. While it is worth remembering that because we don't know if fragments are homologous does not mean that they are not, this issue requires investigation. The debate seems to reflect the 'purist' vs 'pragmatist' approach. On that note I would like to share a comment by Adrian Gibbs, developer of the RAPDistance Programs, who told me that using RFLPs for phylogenetic inferences might be considered by some as following a "suck it and see" approach with the result determining its worth. It remains to be determined precisely what is the 'margin of error' for determining the true relationships among bacteria that is inherent in RFLP and other molecular typing methods that sample only a portion of the genome. And it remains to be determined in each instance what the significance of such error is to our understanding of epidemiology or population structure. For example, subtle differences in the genome of strains may identify differences that are not epidemiologically significant. Until sequencing complete genomes is inexpensive, rapid and commonplace the use of a polyphasic approach to identification will help to resolve some of the uncertainty. Indeed in the last 10 years using multiple molecular typing systems that complement each other has become increasingly more common. Despite theoretical criticisms, PFGE (and other RFLP data) are currently used extensively (and often in conjunction with other typing methods) to make phylogenetic inferences that are improving our understanding of relatedness among bacteria, viruses, and fungi.

PFGE and H. influenzae. CHEF PFGE, now a well-established methodology in many pathology laboratories, is most often employed to analyse relatively small sets of

isolates related to outbreaks of disease. Its use in the analysis of *H. influenzae* began in the early 1990s when it was used to construct a physical map of Hib strain Eagan and to characterise ten genetically heterogenous Hib isolates previously typed by MLEE (Butler and Moxon, 1990). It has also been used to study insertion mutations in the transferrin binding system of Hib (Curran *et al.*, 1994) and other types of PFGE have been used to estimate its genome size (Kauc, 1992). Results of PFGE of 29 epidemiologically unrelated Hib isolates demonstrated evolutionary divergence consistent with that of multilocus enzyme electrophoresis (Arbeit *et al.*, 1990). Until the last several years in which a number of studies have been completed, PFGE, had not been applied to the study of a large number of clinical isolates of Hib in population genetic studies. In conjunction with mathematical analysis the RFLPs generated by PFGE in this study were used to provide numerical estimates of genetic diversity among 213 isolates to construct a dendrogram and to assess the correlation of genetic types with the epidemiology of Hib.

1.8.3.4 Polymerase chain reaction based methods

The polymerase chain reaction (PCR) is an *in vitro* DNA procedure that can, in a matter of hours, isolate and amplify a specific segment of DNA by as much 10^8-fold. The overall ease of performance and the speed of PCR methods are well established and PCR is commonly used in the genetic typing of bacteria. The components are conceptually simple: target DNA to be amplified, oligonucleotide primers, temperature-stable DNA polymerase, nucleotides, and appropriate buffer and salts. The power of PCR is in its theoretical twofold amplification during each cycle, a cycle being defined by temperature shifts in which DNA is serially denatured, annealed to primers, and enzymatically extended. Despite these operational simplicities, users must contend with unique problems of optimising individual assays (e.g., determining optimal primer design and temperature conditions). Since contamination with even one copy of target DNA can produce a false positive result, some workers recommend physically separate work areas for preparing the samples and for performing the reactions. Speed, lower costs, the need for only a small sample, and a protocol with fewer steps are advantages offered by PCR. While PCR has proven to be an extremely important typing tool, one must be aware that the level of reproducibility and the discriminatory ability vary considerably among PCR methods.

The various kinds of PCR-based methods now used to discriminate bacteria are too numerous to mention in entirety here. Three that have been used in the study of *H. influenzae* are described, namely, PCR-RFLP, RAPD/AP-PCR and ERIC.

Polymerase chain reaction-restriction fragment length polymorphism. Polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) analysis involves amplifying a known sequence, cutting it with restriction enzymes, and comparing restriction fragments of the amplified DNA from different strains. It offers advantages over the conventional DNA restriction analyses because no southern blotting or probing is necessary and problems with DNA base modification (methylation) are not encountered. Virulence genes and ribosomal operons have been targeted for PCR-RFLP subtyping.

Randomly amplified polymorphic DNA-PCR. Randomly amplified polymorphic DNA-PCR (RAPD-PCR) (also referred to as arbitrarily primed PCR analysis (AP-PCR)) is similar to RFLP analysis. The process involves randomly amplifying a segment of the target DNA by using a single short primer (~6-10 bp) that does not have any known homology to the target sequence. It assays DNA sequence variation in short regions, but instead of analysing restriction endonuclease recognition sequences, it focuses on PCR priming regions. RAPD analyses are technically simple and often detect variation among isolates that are invariant with conventional RFLP analysis. Criticisms focus on reproducibility problems and incomplete comparison with the powers of other techniques.

Enterobacterial repetitive intergenic consensus-PCR. Enterobacterial repetitive intergenic consensus (ERIC) sequences are conserved regions of DNA dispersed throughout the genomes of Gram-negative enteric bacteria (Pettigrew *et al.*, 2002). Such sequences are typically present at many sites around the bacterial chromosome; when two sequences are located near enough to each other (within a few kilobases) the DNA fragment (called the inter-repeat fragment) between those sites is amplified. The distribution of ERIC sequences varies between strains and ERIC-specific primers have been used to produce fingerprints of bacterial genomes based on the size and numbers of the inter-repeat fragments. The technique appears to have excellent reproducibility and moderate discriminatory power (Arbeit, 1995). A number of PCR techniques

utilising interspersed repetitive elements as primers for the amplification of regions between neighbouring repetitive elements is collectively called rep-PCR.

1.8.3.5 Ribotyping/Southern hybridization

The evolutionary conservation of ribosomal RNA makes it applicable as a universal bacterial probe. Ribotyping or ribosomal DNA RFLP is a method that probes restriction fragments of genomic DNA with cloned probes that contain all or part of the 16S and 23S rRNA genes. It is the most versatile and the most widely used strategy of Southern blot analysis of bacterial genome polymorphism. Patterns are easy to interpret because of a limited number of hybridised fragments (usually 7-12), and the ribotypes are stable and reproducible after subculture of the same organism. Genes coding for rRNA are highly conserved and most bacteria have multiple ribosomal operons. Thus, probing permits inter-species and intra-species discrimination (Stull *et al.*, 1988). The method can be used for long-term epidemiological studies since the ribotype is stable. Its discriminatory power is lower compared to PFGE (Busch and Nitschko, 1999), and at a level equal to or inferior to that of MLEE (Struelens, 1998). This is related to the fact that ribosomal operons cover less than 0.1% of chromosomal length and tend to cluster in one particular region of the genome (Struelens, 1998). Ribotyping has been shown to easily distinguish Hib isolates, however, it cannot be used to reliably cluster apparently related patterns because differences in any two ribotypes, despite being stable, may not be representative of the majority of the genome; indeed, rRNA operons are clustered and cannot be considered to be a random representation of the bacterial genome (Leaves and Jordens, 1994).

1.8.3.6 Sequence typing

Ultimately all molecular genetic methods for distinguishing organisms are based on differences in the DNA sequence. Logically, then, DNA sequencing would appear to be the best approach to differentiating subtypes. But, while DNA sequencing remains the most sensitive method of detecting polymorphism within a gene, there are practical constraints on its use. The major limitation to its use is the size constraints of the DNA that can be practically sequenced. A very small region of the chromosome should be targeted for sequencing and this region must meet several structural criteria (e.g., it must consist of a variable sequence flanked by highly conserved regions) before it can be used for strain differentiation. Further, it should not be horizontally transmissible to other strains of the species. Unfortunately for bacteria, few sequences meet these criteria

(Olive and Bean, 1999). Thus, the development of sequence typing hinges on identifying appropriate genomic targets of interest and determining their sequence variability (Yakubu *et al.*, 1999). To optimise the use of genomic sequencing within and beyond the research laboratory, simplified, automated sequencing methods need to be more generally available.

DNA sequencing of a particular locus. By using PCR to amplify a known DNA segment and automated techniques to sequence the PCR product, it is now feasible to compare multiple isolates by sequencing each one at the same locus. In most cases for diagnostic purposes, the following genomic structures are chosen as a target: 1) DNA sequences bearing the code for toxic or pathogenic factors, 2) DNA sequences of specific antigens, 3) specific DNA plasmid sequences, 4) DNA sequences bearing rRNA codes, and 5) small sequences, mostly species specific, that are non-coding (Harmsen *et al.*, 2001).

Since the early 1980s molecular techniques have been used to determine the nucleotide sequence of ribosomal RNA (rRNA) from various prokaryotes. According to 16S rRNA analyses, life on earth has been classified into three domains: the Bacteria, the Archaea, and the Eucarya (Woese, 1987). Because certain ribosomal sequences are highly conserved across the bacterial kingdom, primers that will amplify ribosomal sequences from essentially any bacteria can be defined. By sequencing the amplified PCR product and analysing the nucleotide sequences at the relatively variable areas within the ribosomal operon, the bacterial family, genus and species can be identified. Since DNA is more easily isolated and manipulated, analysis of 16S rRNA gene sequence data (i.e., rDNA) has subsequently become the most well used method of single locus typing.

Single locus typing has been applied to analyses of bacterial populations (Dubose *et al.*, 1988) (van Loo *et al.*, 1994), epidemiology studies (Swanson *et al.*, 1998) and outbreak situations (Hoe *et al.*, 1999).

Multilocus sequence typing. Multilocus sequence typing (MLST) is based on the well tested principles of multilocus electrophoresis, but it assigns alleles at each site directly by nucleotide sequencing rather than indirectly from electrophoretic mobilities of their gene products in starch gels. It is based on direct sequencing of *c.* 470-bp PCR products of six housekeeping genes. Strain associations shown by MLST were consistent with

clonal groups previously established by MLEE (Maiden *et al.*, 1998). An important advantage of MLST over other typing methods, such as RFLP, RAPD, and PFGE is that the sequence data are truly comparable between laboratories (Maiden *et al.*, 1998). It permits comparison of results from different laboratories without the need to exchange or to take variations in electrophoretic techniques into account. Therefore, it is well suited to the construction of global databanks. Further, MLST is more amenable to quantitative analyses allowing the establishment of quantitative genetic relationships between isolates. Finally, MLST has proven to be especially suitable for studying longer-term and global epidemiology (van Loo *et al.*, 1994).

Although MLST could well become the 'gold standard' method for study of the long-term evolution and global epidemiology of bacteria, it is less sensitive to microevolution than PFGE which means that typing methods such as PFGE that are very sensitive to microevolution will continue to be used to reveal minor differences between related strains (Yakubu *et al.*, 1999).

1.9 Genetic Diversity of Haemophilus influenzae

It had been known for some time prior to the advent of molecular technology that there is diversity within the species of *Haemophilus influenzae*. Early on after its discovery it was noted that some strains of *H. influenzae* were encapsulated and others were not though it was some 30 years before Margaret Pittman's seminal work describing the six capsular serotypes was completed in the early 1930s (Pittman, 1931). It was another 40 years before Mogen Kilian's extensive taxonomic study clearly demonstrated the heterogeneity of *H. influenzae* that had been repeatedly suggested by strain specific differences in pathogenicity and various biochemical properties (Kilian, 1976). Kilian and several others noted a correlation between biotype and the source of *H. influenzae* isolates and it was reported that both antibiotic resistance and type b capsule production were associated with biotype I and II organisms (Albritton *et al.*, 1978). While revealing diversity among *H. influenzae* isolates this data provided limited insight into its scope and significance. Thus, by the late 1970s our understanding of the diversity of *H. influenzae* was still primitive with the body of research based on morphological, biochemical, serological and physiological characteristics. The next 25 years, however, were incredibly fruitful. A fundamental understanding of the population structure and diversity of bacteria, including *H. influenzae* and, in particular, Hib was acquired. This

was due primarily to developments in detecting proteins and nucleic acids. The data that has been acquired regarding the genetic diversity of *H. influenzae* and Hib must be considered in the context of what is known about the extent of genetic variability within the so called bacterial species and the structure of bacterial populations.

1.9.1 Bacterial population structure

The genetic structure of microbial populations has been the subject of growing interest and considerable activity since the early 1980s (after being largely ignored during the burgeoning of descriptive population genetics in natural populations during the 1960s and 1970s). The development of bacterial genetic population studies truly emerged when Selander and colleagues applied multilocus enzyme electrophoresis to the study of bacteria and our present knowledge of the bacterial population structure largely stems from work they published in the 1980s (Selander *et al.*, 1987).

Population genetics is concerned with genetic variation (basically within and between species), its amount in nature, its distribution over time and space, the forces that affect it and its biological significance -- the raw stuff of evolution (Young, 1989). In other words, it can be considered to be the study of natural bacterial genetic diversity arising from evolutionary processes. It is now widely understood that the findings of population genetics research have important implications for understanding the origins and spread of microbial disease and can be useful in advancing our understanding of pathogenesis and epidemiology and in designing effective public health measures (Musser, 1996). Before continuing the discussion on bacterial population structure in section 1.9.1.3 it is appropriate to review the current understanding of the terms 'species' and 'clone' as they apply to bacterial populations.

1.9.1.1 The species

The biologic definition of species (traditionally the most important taxonomic category in biology) when applied to bacteria remains consistently controversial and is undergoing continuous refinement (van Belkum *et al.*, 2001). The basic issue is whether bacteria exist as a series of asexual clones or as promiscuous, freely intermixing populations. For more than 50 years the discussion on species definition has been dominated by that of Mayr and others who define species as "groups of actually or

potentially interbreeding natural populations, which are reproductively isolated from other such groups" (Pennington, 1994). It is generally accepted that a species is a population whose members freely and under natural conditions exchange genes with one another and that the members of a species cannot exchange genes with members of other species. This concept is, of course, only applicable to sexually reproducing organisms. Prokaryotes, which reproduce by binary fission, are asexual (at least by the standard definition of sexual reproduction). Therein lies one problem in assigning the classic definition of species to bacteria.

A second problem is that in contradiction to the asexuality of binary fission, bacteria possess a range of mechanisms for genetic exchange and rearrangement, which include conjugation, transduction and transformation, plasmids, phages and transposable elements and seem to share a wide gene pool that reaches beyond the species level. The apparent promiscuity of certain genetic elements begs the important questions as to whether bacterial populations are composed of reproductively isolated species and whether we should attempt to apply the traditional species concept to bacteria or should consider an alternative image for defining bacterial species. Maynard Smith posited whether is it better to take a wholly gene-centred view and consider the bacterial cell - or rather its chromosome - as merely an alliance of genes, analogous to a football team, composed of players from many different countries, all liable to be to transferred at any time (Maynard Smith, 1990).

In practice most systematists eschew reproductive isolation as a criterion and make inferences based on morphological, biochemical, serological, and cytogenic features (Pennington, 1994). Subjective judgment is certainly required. Consider, for example, that in Bergey's Manual Staley and Krieg define a bacterial species as a collection of strains that share many features in common and differ considerably from other strains. They further state that a species consists of the type strain and all other strains that are considered to be sufficiently similar to it as to warrant inclusion with it in the species (Staley and Krieg, 1984). The imprecision of 'differ considerably' and 'sufficiently similar' represent the state of the dilemma.

Sneath's ground breaking work in the 1950s in numerical taxonomy methods has been applied to quantitative analysis of traditionally phenotypic characteristics and it is now generally accepted that for an isolate to belong to a given species there must be at least

80-85% similarity among independent covariate characteristics (Janda and Abbott, 2002). Sneath has remarked, however, that "good scientific judgment in the light of other knowledge is indispensable for interpreting the results of numerical taxonomy" (Pennington, 1994).

The use of DNA hybridisation was the first major leap in the use of molecular technology to define species. An ad hoc committee of the International Committee for Systematic Bacteriology concluded that the phylogenetic definition of a species would generally include those strains with 70% or greater DNA-DNA homology at a thermal stability within 4°C of that of homologous reassociated DNA (Wayne *et al.*, 1987). Consequently, species are defined on the basis of a primitive form of genetic typing: resemblance between genomes is arbitrarily characterised on the basis of a technology that superficially scans for genome identity (van Belkum *et al.*, 2001).

16S rDNA sequencing (sequencing of the DNA that encodes the 16S RNA subunit) is the single most common molecular technique presently being used for bacterial species identification. While there are no consensus guidelines for its use, it is generally accepted that an unidentified isolate whose rDNA is <97% similar to those of the isolate's closest phylogenetic neighbours constitutes a new taxon (Janda and Abbott, 2002). Despite its widespread and prolific use several concerns cloud the use of rDNA sequencing data for taxonomic relatedness. The lack of consensus on the cutoff point for relatedness (one study suggests that it should be >99% similarity) is one such concern. Another is that 16S rDNA relatedness may not be recognised by other genetic tests (three *Edwardsiella* species recognised as the same species by 16S rDNA with only 0.19 to 0.65% divergence that were genetically distinct by DNA hybridisation and ecological niche has been reported) (Janda and Abbott, 2002). This is complicated by significant sequence variation among identical prokaryote species with multiple rDNA sequences deposited in public sequence databases (Kolbert and Persing, 1999).

An alternative but more generally phylogenetic concept proposed by Cracraft defines a species as a "group with a common origin that is composed of the smallest diagnosable cluster of individual organisms within which there is a parental pattern of ancestry and descent" (van Belkum *et al.*, 2001). It has been proposed that a microbial species should correspond to a discrete typing unit to be valid, in which case the species definition is

clear cut; or, on the other hand, species can be considered "condensed nodes" in an "otherwise cloudy, confluent taxonomic space" (van Belkum *et al.*, 2001).

Dykhuizen and Green's proposal that gene trees can be used to create an operational definition of bacterial species is an attempt to define species from a biological approach (in contrast to the phenetic or phylogenetic approaches mentioned above); some of the difficulties and complications of this approach are discussed in their 1991 publication (Dykhuizen and Green, 1991).

Current definitions, whether conceptualised by taxonomic, phylogenetic or biologic approaches, allow that the concept of species involves making subjective judgement and notably leave precise species boundaries unclear while simultaneous application of varied definitions is inconsistent and confusing. Perhaps now that bacterial population studies are expanding and variation in large samples of the population can be measured at multiple loci directly by nucleotide sequence determination further insight into what constitutes a bacterial species or even whether such an entity exists will be forthcoming. For now, because the term cannot be avoided the word 'species' is probably most often used to refer to assemblages of isolates defined as species by traditional microbiological criteria (Spratt and Maiden, 1999).

1.9.1.2 The clone

It has been noted that identification of clones must be based on thorough monitoring of several (molecular) markers of sufficient discriminatory power and that the definition of "sufficient" is sometimes the subject of intense debate (van Belkum *et al.*, 2001). Thus, the issue of precise boundaries also surrounds the definition of clone but there is a reasonable consensus on a working definition.

The term was first used by Herbert John Weber in 1903 to designate a population in which all members have been derived from one and the same progenitor by non-sexual multiplication (Orskov and Orskov, 1983). In 1982 at a time when significant population studies were just beginning to emerge, investigators at an international workshop entitled "The Clone Concept in Epidemiology, Taxonomy, and Evolution of the Enterobacteriaceae and other Bacteria" agreed that "for the present the word clone will be used to denote bacterial cultures isolated independently from different sources,

in different locations, and perhaps at different times, but showing so many identical phenotypic and genetic traits that the most likely explanation for this identity is a common origin" (Orskov and Orskov, 1983). Clearly a more precise definition was needed.

Milkman, noted for his pioneering work on bacterial population studies, defined a bacterial clone as a group of organisms or cells descended without recombination from a single ancestor (Milkman and Stolzfus, 1988). This is now the generally accepted definition of clone and it may be assumed that members of a clone are genetically identical descendants of a single ancestor whereby all evolutionary change through time is the result of the mutational process or other non-recombinant processes. In contrast, non-clonal or sexual evolution occurs by the horizontal transfer of genetic material between unrelated individuals through the process of recombination (Guttman and Dykhuizen, 1994).

Spratt and Maiden's assertion that "asexual bacterial populations exist as assemblages of independent lineages (i.e., distinct clonal lineages) which are stable, with evolutionary change occurring only by *de novo* mutation, although diversity may also arise by other mechanisms, e.g., the gain and loss of plasmids or the movement of insertion sequences" further clarifies the current definition popularly applied to the nature of bacterial clones (Spratt and Maiden, 1999).

1.9.1.3 Clonality in bacterial populations

The fact that, in some cases, bacterial pathogens represent widespread members of a single clone was first observed by Duguid and collaborators in the 1960s. They were struck by the uniform properties of so-called FIRN *Salmonella typhimurium* bacteria, which are non-fimbriate and cannot ferment rhamnose or inositol. They showed that three independent mutations leading to lack of piliation and an inability to ferment rhamnose were conserved in independent bacteria isolated worldwide and that the original strain had diversified sufficiently for 27 phage types and 22 biotypes to be recognised among FIRN isolates (Morgenroth and Duguid, 1968).

In the mid-1970s the Orskovs concluded that enterotoxigenic strains of *Escherichia coli* represented a few distinct clones because of the association of diarrhoeal disease with

particular O:H serotypes that were otherwise exceedingly rare among the 20,000 *E. coli* isolates they had examined (Orskov *et al.*, 1976). By the early 1980s the clonal population structure of *E. coli* was demonstrated genetically by multilocus enzyme electrophoresis (MLEE) studies (Selander and Levin, 1980; Whittam *et al.*, 1983). The introduction of MLEE to the study of bacterial populations marks a turning point from which bacterial population genetics emerged as an area of considerable activity and interest.

The most extensively studied bacterial population is that of *Escherichia coli* from which our present knowledge on the dynamics of the genetic structure of bacterial populations stems (Milkman, 1973; Selander and Levin, 1980; Whittam *et al.*, 1983; Hartl and Dykhuizen, 1984; Selander *et al.*, 1986). *E. coli* contains thousands of electrophoretic types (ETs). The repeated isolation of certain ETs over decades and continents plus statistical analyses of the genetic distances between ETs led to the concept that *E. coli* is clonal.

Haemophilus and *Neisseria* species have also been important in the early studies on bacterial populations (Musser *et al.*, 1985; Olyhoek *et al.*, 1987; Caugant *et al.*, 1987a; Caugant *et al.*, 1987b; Caugant *et al.*, 1988; Musser *et al.*, 1988). Data collected using MLEE showed that the amount of diversity that exists between independent isolates within species was less than would be expected if all alleles were randomly distributed. Thus, a non-random association of alleles (linkage disequilibrium) was characteristic of bacterial populations and it was concluded that many species of bacteria were clonal in structure. One mechanism that accounts for this structure is periodic selection, in which a rare generally favourable mutant allele rises in frequency, carrying the rest of the genome along as a passive hitchhiker (Milkman and Stolzfus, 1988).

We need to remember that although bacteria are largely asexual, they possess a number of mechanisms for the horizontal exchange of chromosomal DNA. The three most important mechanisms, conjugation, transduction and transformation are properly called parasexual processes as none of them involves wholesale exchange of chromosomal genetic material. Rather, they permit the transfer of fragments of chromosomal DNA, and horizontal genetic exchanges lead to bacterial genomes being pocked by small chromosomal replacements from other lineages, a process that has been referred to as localised sex (Maynard Smith *et al.*, 1991). Plasmids, prophages, transposons and

insertion sequences can also be transferred horizontally, providing mechanisms for mobilising DNA among distantly related bacteria.

A clonal structure implies that recombination is a relatively rare event in these populations but it does not prove that members of a particular clone are identical at all gene loci (Maynard Smith, 1990). To explain the contradiction between clonality and recombination the idea of *clonal segments* (i.e., a chromosomal region not yet broken up by recombination) was introduced in which the bacterial chromosome is the stable *clonal frame* derived from a single common ancestor within which individual segments are inserted from a variety of sources (Milkman and Stolzfus, 1988; Maynard Smith *et al.*, 1991). This mosaic structure has been demonstrated in a number of genes but is not universal (Maynard Smith, 1993a) and is limited by the degree of homology between recombining DNA molecules. It has been estimated that homologous recombination requires as little as 70% nucleotide sequence identity between the recipient and donor bacteria. A consequence of this promiscuity is that bacteria can recruit variation from other members of the same species and, at lower frequencies from related species.

As it is unlinked to reproduction, the frequency of recombination could vary from very low (or zero) to very high. Localised sex (i.e., small chromosomal replacements from other lineages) can disrupt clonal population structure enabling mutations to escape the lineage in which they arose, thus, moderating the power of diversity-reducing events to purge the variation (through periodic selection) in a bacterial population. In other circumstances horizontal genetic exchange can be a cohesive evolutionary force, by enabling a variant allele to spread horizontally in a population, replacing other variants and reducing diversity (Spratt and Maiden, 1999). The extent and impact of recombination versus mutation on the diversification of bacterial clones is not clear but recent evidence suggests that in many species recombination contributes more greatly to clonal diversification than do point mutations (Spratt *et al.*, 2001).

Maynard Smith *et al.* (1993b) statistically tested the extent of clonality and identified four types of bacterial populations. They argued 1) that only certain species (*Escherichia coli, Salmonella spp., and Haemophilus influenzae*) have clonal population structures where little recombination has occurred in the evolutionary history of the species, 2) that others (*Neisseria gonorrhoeae*) are panmictic or fully sexual with random association between loci owing to extensive genetic recombination, 3) that still

others (*Neisseria meningitidis*) only appear to be clonal owing to rapid epidemic spread of panmictic bacteria and, 4) that other *Rhizobium*-like populations are panmictic at the fine scale but do not recombine between populations (Maynard Smith *et al.*, 1993b). Over the short term, all these bacteria have a clonal population structure which can be obscured by recombination or amplified by spread, depending on the species and in some cases even on the specific clonal grouping; thus, the distinction between clonal, panmictic and epidemic populations structures reflects different ratios of recombination to spread (Achtman, 1994).

It is thought that the extremes of strictly clonal and non-clonal populations structures are rarely found in bacterial species, and that most bacterial populations occupy a middle ground where recombination is highly significant in the evolution of the population, but is not sufficiently frequent to prevent the emergence of clonal lineages (Spratt and Maiden, 1999). Further it has become increasingly clear that bacterial populations are complex and may not always be adequately described as being highy clonal, weakly clonal or non-clonal; for example, populations may be non-clonal in the longer term, although the emergence of unstable clones may be an important feature of the population (Spratt and Maiden, 1999). An important fact to consider is that bacterial populations do not seem to evolve toward a "single most fit clone" occupancy of a given niche. Even in a simple environment such as a stationary batch culture, bacteria appear to maintain an interactive polymorphic state (van Belkum *et al.*, 2001).

1.9.2 Genetic diversity in H. influenzae populations

As noted previously by the late 1970s our understanding of the diversity of *H. influenzae* was still primitive with the body of research based on morphological, biochemical, serological and physiological characteristics. Since then many different typing schemes and hundreds of studies have added to a growing knowledge of the population structure of Hib. For example, besides biotyping, OMP subtyping, and MLEE, schemes used for subtyping Hib include plasmid analysis (Mendelman *et al.*, 1985), whole cell polypeptides (Pennington and Freebairn, 1989), DNA fingerprinting (Mencarelli *et al.*, 1993), ribotyping (Leaves and Jordens, 1994), *cap* region RFLP patterns obtained by digestion with *Eco*RI (Musser *et al.*, 1988), DNA sequencing and PCR fingerprinting (Falla *et al.*, 1994) and long PCR ribotyping (Smith-Vaughan *et al.*,

1998). Recently, RFLP analysis using PFGE has been used to analyse small samples of Hib (Tarasi *et al.*, 1998; Mitsuda *et al.*, 1999; Saito *et al.*, 1999). In this section I will discuss OMP typing, the first method to provide data that seriously identified Hib diversity and MLEE, the method used to provide a population genetic framework for the analysis of variation in other characters and which was used in a seminal piece of work that revealed the clonal population structure of Hib.

In the early 1980s OMP subtyping was found to be a useful technique for investigating the epidemiology of Hib and substantial gains were made in understanding its diversity (Barenkamp *et al.*, 1981; Loeb *et al.*, 1981; Granoff *et al.*, 1982). The first report using OMP data in an epidemiological study provided information on the evolution of *H. influenzae* strains among which the presence of six serotypes itself suggests evolutionary divergence. This was supported by the lack of relatedness of Hib OMP patterns (8 were found among 28 isolates) with those of four other capsular serotypes (Loeb and Smith, 1980). In addition, considerable variation among NTHi strains left open the question of whether they evolved before encapsulated forms and subsequently gave rise to the various serotypes and/or whether they evolved from any of the serotypes (Loeb and Smith, 1980).

The limited number of OMP subtypes found in initial studies suggested that a small number of distinctive clones were responsible for most Hib disease in the United States. Barenkamp, et al. described 9 subtypes among 50 isolates from patients at St. Louis Children's Hospital with just 5 types accounting for 92% of the strains (Barenkamp *et al.*, 1981).[2] The designation and proportion of these 5 types being: 1L (14%), 1H (31%), 2L (22%), 2H (10%), 3L (16%). The remaining isolates had unique profiles different from the 5 predominant types and from each other. Subtype concordance was observed among epidemiologically related strains. Further study identified a total of 21 distinctive OMP subtypes among isolates from 256 patients hospitalised in 22 states; six of these subtypes accounted for 89% of the isolates and three of these – 1H, 2L, and 3L – were prevalent in all the tested regions (Granoff *et al.*, 1982). Strains with the 1H OMP pattern were most common representing 37% of all endemic cases and associated

[2]In the "Barenkamp" system for OMP classification subtypes are designated 1, 2, 3 and so on, according to the pattern produced by proteins with a molecular weight of 25,000-40,000. An H (heavy) or L (light) category is assigned based on the apparent molecular weight of a heat-modifiable protein of 50 kilodaltons (kDa) or 49 kDa, respectively. A few strains with a heat-modifiable protein different from the H or L protein are designated U or unclassified .

with a tenfold increase in the secondary spread of disease in day-care centers that was not explained by increased transmissibility or by enhanced ability to colonise the nasopharynx (Barenkamp *et al.*, 1981).

Compared to the findings in the United States, a remarkable homogeneity among 80 randomly selected Hib isolates from patients with meningitis was seen in the Netherlands showing that they were not strongly subject to genetic change (van Alphen *et al.*, 1983).[3] 67 of the 80 strains studied had identical OMP patterns designated subtype 1. Among the 13 other isolates 4 different patterns were seen. It was further revealed that 75% of the isolates were characterised by a particular OMP pattern (subtype 1), LPS serotype (type 1), and biotype (type I). The uniformity of the strains indicates that they are not strongly subject to genetic variation. The lack of uniformity among the isolates in the United States is likely caused by the distribution of various subtypes (among which is one that is predominant in the Netherlands) and is not simply a matter of geographic area because subtypes in the United States were nearly randomly distributed around the country (van Alphen *et al.*, 1983).

Examination of strains from other Western European countries (Denmark, Sweden, Finland, Iceland, Norway, France, Germany and the United Kingdom) showed that only a few subtypes of Hib predominated and that the distribution of these subtypes showed distinct geographic differences (van Alphen *et al.*, 1987). The predominant subtype (OMP 1, LPS 1, Biotype I) in the Netherlands was predominant in Western Europe except for Iceland where it was almost absent. A new type (OMP 2, LPS 9, Biotype I) not previously described comprised 83% of the isolates in Iceland. It was not found in other countries and was not found to be associated with a particular disease type. van Alphen's OMP subtype 2 is identical to Barenkamp's 2L but several American isolates of this subtype that were tested were not LPS serotype 9. The uniqueness of the "Iceland" subtype is difficult to explain. It may be that this type is present but was not found in the countries from which only a few isolates were studied.

---

[3]The system used by van Alphen for OMP classification assigns isolates to subtypes based on the position of the major outer membrane proteins, a, c, d, and e described by Loeb (Loeb *et al.*, 1981). The predominant pattern was designated subtype 1. Patterns similar to subtype 1 that showed a difference in protein d were designated 1a and 1b. Two other patterns were designated subtype 2 and 3 based on differences in the mobility of proteins c and a, respectively. Comparison with Barenkamp's nomenclature revealed subtype 1 was identical to 3L, subtype 1a was unique, subtype 1b resembled 16L, subtype 2 was identical to 2L, and subtype 3 was identical to 6U.

Although the majority of infections worldwide are caused by just a few OMP subtypes (Musser *et al.*, 1990), the predominant subtypes differ in different parts of the world. The subtypes found in the Far East (Weinberg *et al.*, 1989), the Gambia (Bijlmer *et al.*, 1992) and among Aboriginal children in Australia (Gilbert and Clements, 1993; Hansman and Lawrence, 1993) differ from those found elsewhere in the world. Four unique OMP types (not previously described by either OMP system) were revealed in the Gambia with one of them accounting for 54% of the isolates; only one type, van Alphen's type 2 that corresponded with Barenkamp's type 2L, had been previously described among European and U.S. isolates (Bijlmer *et al.*, 1992).

In Australia, 76% of 59 strains belonged to a single OMP subtype equivalent to subtype 3L in the Barenkamp scheme; the remaining type b strains belonged to five other OMP subtypes (Hansman and Lawrence, 1993). The distribution of OMP subtypes from rural Aboriginal children and Caucasian children in Melbourne and Sydney differs. In the latter group 183 of 220 isolates were the same subtype and identical with the predominant European subtype (1VA in Barenkamp's scheme); 3 additional isolates were of the almost identical subtype 3L (Barenkamp scheme) (Clements *et al.*, 1992). Together 1VA and 3L represented 85% of isolates. By contrast only 13% of 55 isolates from aboriginal children were type 1VA and none was subtype 3L. The predominant subtype was an apparently unique local subtype, designated 1NT, that was almost identical with subtype 1L (Barenkamp scheme) (Gilbert and Clements, 1993). It accounted for 82% of isolates from Aboriginal children compared with only 1% of isolates from Melbourne and Sydney. Ten other unique subtypes were found among the urban and rural isolates. The marked discrepancy between OMP subtypes causing Hib disease in two major groups of children is intriguing and should be confirmed by testing isolates from Caucasian children in Western Australia or the Northern Territory

In Melbourne the distribution of subtypes also varied in different age groups suggesting that less common subtypes of Hib are more likely to cause disease in younger children (<12 months old) in which the incidence of Hib disease is higher because of increased susceptibility or exposure (or both) (Gilbert and Clements, 1993).

Takala *et al.* (1987) reported that in Finland type b strains of OMP subtypes 1 and 1c (3L and 1L, respectively in the Barenkamp scheme) are non-randomly associated with different types of invasive disease. The observation that strains of subtype 1c caused proportionally more meningitis and less epiglottitis than did those of subtype 1 was

interpreted as evidence of a true difference in virulence between isolates expressing those subtypes. It was hypothesised that the subtype 1c protein marks a clone with special virulence properties. To date no evidence has confirmed this but the results of OMP subtyping of 587 world wide Hib isolates were consistent with the interpretation that in Finland the type 1c OMP subtype frequently marks a subclone that is especially successful in invading the CSF (Musser *et al.*, 1990).

In the 1980s multilocus enzyme electrophoresis was first used to study 177 isolates of Hib, largely from the United States, and a basically clonal structure of Hib populations was revealed and they were marked by temporal and geographic variations (Musser *et al.*, 1985). The use of MLEE in an extensive survey reported in 1990 of 2209 encapsulated *H. influenzae* isolates from all over the world revealed that there are two primary phylogenetic divisions, designated I and II, within the encapsulated *H. influenzae* and that the population structure is tightly clonal. It was revealed that capsulated *H. influenzae* belong to twelve major lineages, designated A-L, each of which belongs to one of the two primary divisions. Division I contains all serotype c and d isolates studied, one lineage of serotype a isolates, and 95% of type b electrophoretic types. Division II contains all serotype f isolates studied, a second lineage of a isolates, and a second group of serotype b isolates. A serotype belongs to only one or a few lineages, and these lineages are not shared between serotypes (Musser *et al.*, 1990).

There is an interesting distinction between the capsule locus of division I and the capsule locus of division II type b isolates. In division I isolates, the IS*1016* element is associated with the capsule locus. In division II isolates, the element is present, but not associated with the capsule locus. IS*1016* is also found at either end of the capsule locus of type a-d strains that belong to division I, although the capsule loci are present as a single copy. Division II type b strains also contain a single copy *cap* locus. These strains are less virulent in an infant rat model of pathogenesis and are rarely responsible for invasive disease cases. Thus, it appears that the 1.2 kb deletion of the bridge region (refer to section 1.5.3.1) is an ancestral mutation that causes a selective pressure for maintenance of the duplicated capsule locus state (Preston and Apicella, 1999).

Most of the invasive disease worldwide was caused by serotype b strains of 9 clones. Strong patterns of geographic variation on an intercontinental scale, in both the extent

of genetic diversity and the clonal composition of populations was noted. Only one situation was identified in which there was a statistically significant association of a clone with a particular syndrome - ET 21.8 in Finland with meningitis - but no other obvious associations of clones and invasive disease were revealed (Musser *et al.*, 1990).

Musser et al's analysis suggests that the present distribution of Hib clones is, in part, related to patterns of racial or ethnic differentiation and historical demographic movements of the human host populations. Briefly, four phylogenetic lineages of clones of serotype b are highly distinctive in multilocus genotype, which means their phylogenetic relationships are relatively deep. Hence, the pattern of global geographic variation in clonal composition of populations that they revealed for Hib does not represent merely a transitional state in the relatively recent spread of new minor mutant variants. The degree of genetic differentiation indicates several of the lineages have long histories of differentiation (nucleotide substitutions have been estimated to accumulate at the rate of 1% per million years); they are old. They proposed two hypotheses to account for this. First, it is suggested that a large component of the geographic variation in clonal composition of Hib seen today reflects an older pattern of differentiation that evolved in relative geographic isolation before the Age of Exploration (beginning about 450 years ago) and has not yet been completely obscured by recent demographic changes. Alternatively, a second hypothesis suggested that there are, in fact, no old human racial or ethnic correlates of the contemporary global pattern of distribution and abundance of the various clones or clone groups. Rather, all the major clonal groups have since earliest times been present in all major geographic regions and human populations. This hypothesis postulates the contemporary pattern of variation in the relative abundance of Hib clones in different geographic regions can be attributed to differential susceptibility to various clones on the part of human populations. A weakness of this hypothesis is that is seems unlikely that all the clones of Hib could have achieved and maintained global distribution while the human species was undergoing geographic differentiation in strong geographic isolation over the past 50,000-100,000 years. For ethical reasons rigorous prospective testing of these two hypotheses is not possible. One critical test comes from an examination of isolates from native or other populations that have experienced very little immigration: hypothesis one explicitly predicts that these populations (aboriginals or others) should have a distinctive set of Hib clones in a relatively high frequency, whereas hypothesis two does not. The bulk of the data favours hypothesis one. Study of isolates from regions of the

world not yet sampled, and especially strains recovered from additional human populations, such as, Australian Aboriginals, that have experienced little recent admixture, will provide critical data bearing on these hypotheses.

Most of what we know about the genetic diversity of Hib in Australia is based on OMP studies that have been shown to be of limited value in detecting genomic variation because only a few OMP subtypes are responsible for most Hib infections. By exploiting the discriminatory power of PFGE we will increase our knowledge of the genetic diversity of Hib in Australia. Because we will be examining a large sample of Aboriginal isolates we will be able to provide some of the 'critical data' bearing on these intriguing hypotheses about the history of Hib clones and how humans may have contributed to their worldwide distribution.

1.10 Aims and Strategies

SIGNIFICANCE OF THE PROJECT

It is not clear what characteristics, if any, of the Hib bacterium are associated with its epidemiology and little is known about its genetic diversity in Australia. The epidemiology of Hib has been studied in Victoria, Sydney, the Northern Territory, the Australian Capital Territory and Western Australia. There are similarities between the main urban centres, especially with respect to the incidence and age-distribution of the most important Hib disease, meningitis. There is greater variability in the epidemiology of the other major Hib disease, epiglottitis. Using the same criteria for diagnosis, three population-based studies in Victoria, Sydney, and Western Australia showed significant differences in the incidence of epiglottitis; in Victoria it was nearly twice that in either Sydney or Western Australia. In a smaller study in the ACT the incidence was similar to that in Victoria. These differences may be explained in part by socioeconomic and/or climatic factors. However, the possibility that they are related to differences in strains of Hib in these geographically separate areas has not been explored.

The other major difference in the epidemiology of Hib disease in these Australian studies is between Aboriginal children and (mainly) non-Aboriginal children in metropolitan centres. The incidence of Hib disease in rural Aboriginal children is several times that in other Australian children, the disease occurs at a significantly lower age and epiglottitis is virtually unknown. Meningitis and pneumonia with bacteraemia

are the commonest manifestations and serotypes other than type b and non-capsulated strains are frequently recovered. This is similar to the epidemiology of Hib disease in other high-risk ethnic minorities and in developing countries. Socioeconomic and genetic factors have been implicated but there are likely to be other factors involved, including distribution of Hib strains in different populations.

Understanding the population structure of pathogenic bacteria may provide insights into the pathogenesis of human infections. Currently, there is no practical typing system for monitoring Hib isolates from vaccine failures, adults and unimmunized children.

GENERAL AIMS OF THE PROJECT

To study the distribution of genetic subtypes of *Haemophilus influenzae* type b among Hib isolates from Australian children (Caucasian and Aboriginal) and adults with invasive Hib disease and from carriers from different parts of the country, to determine whether particular genetic types are associated with different epidemiological and/or virulence factors and to establish methods of typing and identifying Hib for the surveillance of Hib disease following the introduction of Hib vaccines in Australia while generating the baseline data against which future distribution patterns of subtypes can be compared that would be necessary to assess the effects of vaccination on the genetic diversity and structures of populations of Hib.

If time allowed this project aimed to investigate the genetic differences, if any, found among Hib isolates that may be related to epidemiological differences. The search will be for those parts of the genome which are invariably associated with the epidemiological characteristic being studied.

SPECIFIC STRATEGIES OF THE PROJECT

Assess the usefulness of pulsed field gel electrophoresis (PFGE) for detecting the genetic diversity of Hib isolates and acquire Hib isolates recovered from both Caucasian and Aboriginal children and adults from a variety of geographic areas in Australia.

Use molecular genetic typing techniques to determine whether differences in the epidemiology of Hib disease in different geographical regions and ethnic groups in Australia are reflected in the distribution of genetic types of Hib by studying isolates from Caucasian and Aboriginal children and adults with invasive Hib disease and/or

upper respiratory tract carriage to answer the following questions and others about the genetic diversity of Hib: What is the degree of genetic diversity among Hib isolates in Australia? Is the clonality, which has been demonstrated for isolates in the USA and Europe characteristic of the isolates of Hib found in Australia? Are there molecular genetic differences among strains of Hib which will distinguish between: 1) invasive and non-invasive isolates, 2) isolates which cause epiglottitis and isolates which cause meningitis, 3) isolates recovered from patients living in different geographic regions, 4) isolates recovered from Aboriginal and non-Aboriginal children

Using molecular genetic techniques establish a broad database of genetic subtypes of Hib for epidemiological surveillance of Hib disease to: 1) study changes in the distribution of genetic types, if any, after introduction of the Hib vaccines, 2) investigate suspected outbreaks of Hib disease, 3) investigate cases of *Haemophilus influenzae* disease in immunized and unimmunized children and, 4) screen non-typeable *Haemophilus influenzae* isolates from children with invasive disease or upper respiratory tract carriage to test the hypothesis that some are capsule deficient Hib strains (and potentially more virulent than other NTHi).

CHAPTER 2

Methods

2.1 Bacterial isolates (Appendices A.1 – A.6)

A collection of 213 isolates of *Haemophilus influenzae* type b recovered from Aboriginal and non-Aboriginal children (*n=208*) and adults (*n=5*) from diverse rural and urban regions of Australia was obtained for this project. One hundred eighty-nine isolates were recovered from patients diagnosed with Hib diseases. Twenty-four isolates were from individuals who did not exhibit any sign of disease at the time the isolates were recovered. No follow-up information is available as to whether these children subsequently developed Hib disease and the isolates are considered to represent strains found among Hib carriers.

Hib isolates recovered before the introduction of the first conjugate Hib vaccine (PRP-D) for use in children 18 months and older in July 1992 comprise the majority of isolates studied. Seventeen isolates recovered in 1993 are included in the collection although by mid-1993 another three conjugate vaccines (PRP-OMP, HbOC and PRP-T) were available for use in children under 18 months and were free to all children under the age of 5 years (Herceg, 1997). The vaccine uptake rates were estimated to be 45-50% in Sydney (McIntyre *et al.*, 1995) and Darwin (Mitchell and Krause, 1995) and exceeding 90% amongst rural Aboriginal children in the Northern Territory (Mitchell and Krause, 1995) by August -September 1993.

The isolates were obtained from established collections through the generosity of the scientists who shared them. Thirty-four isolates from the Canberra region and 37 from Sydney were obtained from the Canberra Hospital, Australian Capital Territory (courtesy of Jan Bell and Dr Peter Collignon); 20 isolates from Melbourne and 10 from the Alice Springs region were obtained from the Royal Children's Hospital in Melbourne, Victoria (courtesy of Professor Gwendolyn L Gilbert and Frances Oppedsidano); 45 isolates from metropolitan Perth and rural areas of Western Australia were obtained from the Princess Margaret Hospital in Perth, Western Australia (courtesy of Professor G L Gilbert and Leoni Walpington); 10 non-Aboriginal isolates from the Alice Springs region were obtained from the Queensland Institute of Medical Research, Brisbane, Queensland (courtesy of Dr Michael Gratten); 35 isolates from the Alice Springs Region originally from Dr M Gratten were obtained from the Westmead Hospital collection (courtesy of Professor G L Gilbert and Helen Fuller); 19 isolates from Bathurst Island were obtained from the Menzies School of Health Research in

Darwin, Northern Territory (courtesy of Dr Amanda Leach) and 2 isolates were obtained from the Townsville General Hospital, Queensland (courtesy of Chris Ashurst-Smith). The geographic origin of the isolates is depicted in Figure 2.1.

Figure 2. 1 Map of Australia showing the geographic locations from which Hib isolates were obtained

The collection used for this project ultimately included 137 isolates from non-Aboriginals and 76 from Aboriginals comprising 95 isolates recovered from the cerebrospinal fluid of patients with meningitis, 37 isolates from the blood of patients with epiglottitis, and 39 from patients with other diagnoses including 13 isolates from the blood of patients for whom no clinical information was available and 24 isolates from healthy carriers. A breakdown of the distribution of the Hib isolates by geographic location, Aboriginality, and disease association is shown in Table 2.1.

The acquisition of previously collected Hib isolates for this study occurred over a period of 3 years. The isolates were not initially collected for a single purpose and the epidemiological data available varies. While it is not known how medical diagnostic models and detection systems used to recover the isolates may have varied, it is unlikely to be significant among the collections that were accessed. There is also no evidence to suggest that there are strains of Hib not readily recovered using standard specimen collection and culture procedures for *Haemophilus influenzae*. Therefore, it is reasonably assumed that the Hib collection for this project represents a spectrum of isolates that can be considered representative of the overall diversity of Hib in Australia.

Table 2. 1 Geographic location and disease association of 213 Hib isolates recovered from Aboriginals and non-Aboriginals before the introduction of vaccines

Group and geographic location	Number of isolates				
	Meningitis	Epiglottitis	Other	Carrier	Total
Aboriginal	23	0	29	24	76
Alice Springs region, Northern Territory	11	0	28[a]	6	45
Bathurst Island, Northern Territory	0	0	1[b]	18	19
Perth, Western Australia	5	0	0	0	5
Rural Western Australia	7	0	0	0	7
Non-Aboriginal	72	37	28	0	136
Canberra, Australian Capital Territory	20	14	0	0	34
Melbourne, Victoria	10	10	0	0	20
Sydney, New South Wales	21	1	15[c]	0	37
Perth, Western Australia	12	8	5[d]	0	25
Rural Western Australia	4	4	1[d]	0	9
Townsville, Queensland	2	0	0	0	2
Alice Springs region, Northern Territory	3	0	7[e]	0	10
Total	95	37	57	24	213

[a]These isolates were associated with pneumonia (*n=9*), gastroenteritis (*n=9*), acute lower respiratory tract infections (*n=4*), cellulitis (*n=1*), febrile disease (*n=1*), conjunctivitis (*n=1*), otitis media (*n=1*), bronchiolitis (*n=1*), and failure to thrive (*n=1*).

[b]The isolate may be associated with otitis media.

[c]Thirteen invasive isolates recovered from blood (diagnoses unknown) and 2 isolates associated with cellulitis.

[d]The diagnoses for these isolates are unknown.

[e]Invasive isolates recovered from blood (diagnoses unknown).

A complete collection of the isolates used in this study is stored at -70°C in the Australian Capital Territory at the Gadi Research Centre, University of Canberra. A partial collection is stored at the Canberra Hospital.

2.1.1 Identification of bacterial isolates

All of the isolates obtained for this study had previously been identified by conventional methods. These included identification by typical Gram stain results, colony morphology on chocolate agar, catalase production, X and V factor requirement, lack of haemolysis when grown on tryptic soy agar plates containing 5% horse blood and 1% IsoVitaleX and biotyping. They were also previously tested for capsular type b antigen by either coagglutination or slide agglutination and were stored at –70°C except for 46 isolates from a collection in Western Australia that had been lyophilised.

Upon receipt, frozen isolates were thawed and aliquots were subcultured to chocolate agar plates that were incubated overnight at 35°C at 5% CO_2. Individual colonies were

then subcultured to a second chocolate agar plate that was incubated under identical conditions. Typical colony types found on chocolate agar were tested for X and V factor requirements (Difco). All the isolates were genotyped to confirm that they were capsular type b or a capsular type not type b or that they did not possess the *cap* locus.

2.2 Pulsed field gel electrophoresis

Isolates were independently digested and electrophoresed at least two times. When comparison of fragment size on different gels was questionable, another gel was run so those fragments could be compared in close proximity. Partial digests and lanes with any smearing due to DNAse contamination were repeated. A six day protocol using a 15 well gel that included 12 isolates and 3 controls is summarised in Table 2.2.

Table 2. 2 Timetable of a six day protocol for PFGE of Hib

Day	Activity	Time Required*
Day 1	Primary isolation of bacterium with overnight incubation	< 1 hour
Day 2	Subculture of bacterium with overnight incubation	< 1 hour
Day 3	Preparation of DNA plugs with overnight incubation	2-3 hours
Day 4	Wash and overnight digestion of DNA plugs	5-6 hours
Day 5	Gel set up and overnight electrophoresis run	< 2 hours
Day 6	Staining and photographing gel	< 2 hours
	Evaluation and analysis of RFLPs	Variable, > 8 hours
(Day 7	Photographing gel after 24 hour destaining wash	< 1 hour)

*Times estimates are based on what can be expected with experienced personnel using the methods described and does not include time required to prepare media and reagents or to clean glassware, etc.

For many bacteria, particularly Gram-positives, a lysis step(s) that can require overnight incubation is needed to break up the cell wall prior to preparing DNA plugs. This was not necessary when working with Hib, a Gram-negative bacterium whose cell wall is disrupted in ESP solution. There are a number of other variations that can be applied to PFGE methods, particularly, reducing incubation times, that will decrease the time requirement. Rapid methods have been described for 3 day (Matushek *et al.*, 1996) and 24 hour (Gautom, 1997; Chang and Chui, 1998) protocols.

2.2.1 DNA preparation

The preparation of chromosomal DNA was performed by a modification of the method described by Bautsch (Bautsch, 1992). DNA was prepared in a 1:1 ratio in an intact 1% agarose (SeaPlaque® low gelling temperature agarose, FMC Bioproducts) matrix using cells harvested from an overnight growth on chocolate agar suspended in solution 21

(Lee and Smith, 1988) to an optical density of 1.8 at 650nm. The DNA-agarose plugs (1x8x19 mm) were lysed in a solution of proteinase K (1mg/ml) in 0.5M EDTA and 1% sodium lauryl sarcosine and incubated overnight at 50°C. The plugs were rinsed for 2 hours in TE (10 mM Tris, 1 mM EDTA, pH 8.0) and then stored in fresh TE at 4°C.

2.2.2 Selection of enzymes

H. influenzae has a base composition of 37 mol% G+C. Therefore, enzymes recognising six-base-pair sequences comprising only G-C bases should produce a small number of distinct fragments. *Sma*I (CCC↓↑GGG) and *Apa*I (G↑GGCC↓C) were used.

2.2.3 Restriction endonuclease digestion

Washing of plugs and digestion of the DNA with endonucleases were performed as described by Inglis *et al.* (1993) (Inglis *et al.*, 1993). *Sma*I (Boehringer Mannheim) and *Apa*I (Boehringer Mannheim) were used according to the manufacturer's instructions. For each analysis, one half of a DNA-agarose plug was digested overnight with 30 units of enzyme at 30°C (*Sma*I) or 37°C (*Apa*I). Smaller concentrations of enzyme resulted in partial digestions. Consistent clean cutting was obtained when 30 or more units of enzyme were used.

2.2.4 PFGE analysis

Separation of DNA fragments using a clamped homogeneous electric field device (CHEF DR™II, Bio-Rad) at 200 volts was carried out in 1% agarose (Type II-A: Medium EEO, Sigma) gels made with .5X TBE (0.045 M Tris-borate, 0.001 M EDTA) in 0.5X TBE buffer at 12-14°C. Pulse parameters for both *Sma*I and *Apa*I digests included a ramp of pulses for 6-8 seconds for 7 hours followed by a ramp of 1-38 seconds for 17 hours. Gels were stained for 30 minutes with ethidium bromide (0.5μg/ml) in distilled water, then destained in distilled water for 1 or more hours and photographed (Polaroid 667) using a UV light source. The polaroid pictures of the RFLP patterns were scanned using an Astra UMAX 1200S Scanner, printed out on a Laser Jet 4000N printer and photocopied for inclusion in this thesis.

2.2.5 Molecular weight standards

The DNA molecular weight standards used were λ phage concatemers (Lambda Ladder PFG Marker, New England Biolabs), *Hind*III phage λ fragments (prepared by digesting phage λ DNA with *Hind*III), and the genomic fragments from Hib strain HS008. This isolate came from the Canberra Hospital collection. It was recovered from a sputum

specimen from a four year old female. It was chosen as a standard during the preliminary studies carried out in the process of setting up the PFGE protocol. The fragment sizes of HS008 were estimated using a plot of log molecular weight of λ phage markers versus distance migrated. The estimated sizes of 15 *Sma*I fragments of HS008 are shown in Table 2.3.

Table 2. 3 Estimated size and schematic diagram of 15 *Sma*I fragments* of Hib strain HS008

Schematic diagram	Fragment	kb
	8.1	416
	8.2	256
	8.3d	209
	8.4d	209
	8.5	147
	8.6	107
	8.7	98.5
	8.8t	81
	8.9t	81
	8.10t	81
	8.11	62
	8.12	57
	8.13	43
	8.14	35
	8.15	33

*Fragments are numbered from 1 (according to size – largest to smallest) including each member of a double or triple. The name of the isolate (in this case, 8) followed by a full stop and then the number of the fragment is used to indicate the marker fragment. 'd' or 't' next to an isolate fragment means it is part of a doublet or triple.

2.2.6 Fragment database for *Haemophilus influenzae* type b

A database of the distinct fragments found among the Hib isolates was compiled for each enzyme as a "reference set" of fragments to use in the interpretation of RFLPs. A total of 79 distinct *Sma*I fragments were found among 213 isolates. Fifty-seven distinct *Apa*I fragments were found among 84 isolates. The presence or absence of each of the different fragments for the appropriate enzyme(s) was visually determined for each isolate in the numerical determination of genetic diversity.

2.2.7 Estimation of genetic diversity between isolates and construction of dendrograms

In the analysis of the Hib isolates described here, RFLPs were determined for the entire genome, except for fragments present in such small amounts that they did not bind sufficient ethidium bromide to be visible, or so small (<6 kb) that they ran off the end of the gel during electrophoresis; such fragments would account for a small proportion of the genome.

The numbers and mobilities of fragments were determined by visual examination of the polaroid photographs of the stained gels. Initially, gel patterns were interpreted by two or more individuals and assignment of fragment size was determined by the use of several size standards, including lambda concatemers and the *Sma*I digest of the chromosome of *S. aureus* ATCC 8325 used by Pattee to define the physical map of *S. aureus*. The complexity of visual examination of the large number of gels in this study was diminished by the use of multiple standards on each gel, namely, lambda concatemers and the *Sma*I digest of Hib isolate HS008 (sized using lambda concatemers and *S. aureus* ATCC 8325 from the average estimate of 10 runs). Each gel was interpreted at different times on multiple occasions to confirm interpretation. As the number of typed isolates grew, interpretation was facilitated by the presence of patterns on gels that were indistinguishable from those in the database because they provided another point of reference. The use of a computerised program designed to interpret and store RFLP data would be helpful, though visual inspection, will always be needed to confirm such results.

Numerical analysis of RFLP patterns was performed by identifying the proportion of fragments shared by pairs of isolates using the method of Nei and Li (Nei and Li, 1979) (also known as the Dice coefficient or coefficient of similarity) and calculated as $F = 2n_{xy}/(n_x + n_y)$, where n_x is the total number of DNA fragments from isolate X, n_y is the total number of fragments from isolate Y, and n_{xy} is the number of fragments identical in the two isolates.

A matrix of F values for all pairs of isolates was constructed and a dendrogram computed using the NJTREE program, i.e., the neighbour-joining method of Saitou and Nei (Saitou and Nei, 1987) and the TDRAW program. The programs used for mathematical calculations and construction of the dendrogram are accessible on the

package, RAPDistance Programs, distributed by J. Armstrong, et al., Research School of Biological Sciences, Australian National University (Armstrong *et al.*, 1994).

The Nei and Li similarity matrix for one sample, the epiglottitis isolates from Western Australia, would not compute with the NJTREE program and Jaccard's algorithm found in the RAPDistance Programs was used to build a matrix from which a dendrogram could be constructed. Jaccard's similarity value is calculated as $F = n_{xy}/(n-00)$, where n_{xy} is the number of fragments identical in the two isolates, n is the total number of possible DNA fragments and 00 is the number of those fragments not present in x or y.

2.3 Other methods

Capsular typing using hybridisation Capsular type was confirmed by hybridisation of Southern blots of *Sma*I DNA fragments separated by PFGE with ^{32}P labelled pU082, a 1.9 kb BamHI/EcoRI portion of the 2.1 kb EcoRI fragment cloned into pBR322, which spans type b specific parts of the *H. influenzae cap* locus (Kroll and Moxon, 1988). Hybridisation products were detected by autoradiography. If no signal was detected, isolates were reprobed with pU038, an 18 kb BamHI fragment from wild type Eagen subcloned from cosmid clone pSKH3 into pBR322, which spans serotype specific and non-serotype specific parts of the *cap* locus (Kroll and Moxon, 1988). Isolates which did not hybridise with pU082 were not included in the study. The probes were originally provided by Professor E. Richard Moxon to Professor G L. Gilbert. Only one isolate was found to be non-encapsulated and excluded from the study.

P2 typing The P2 gene of *H. influenzae* encodes the outer membrane protein P2 that accounts for the major protein component of the outer membrane and functions as a porin protein. Hib strains have highly conserved P2 sequences (Munson *et al.*, 1989). A P2 gene probe previously prepared by cloning a 7 kb fragment in the EcoR1 site of pUC9 was obtained from the Royal Melbourne Hospital and used to screen 84 isolates (34 from Canberra and 60 from the Alice Springs region) for polymorphisms (if any) that would be detectable by hybridisation of *Sma*I DNA fragments separated by PFGE.

Southern blotting and hybridisation DNA fragments separated by PFGE were acid depurinated in 0.25 M HCl for 2 x 8 minutes followed by 2 x 8 minutes rinses in 0.4 M NaOH and then transferred onto Hybond-N nylon membranes (Amersham) in 10 mM

NaOH overnight. The membranes were prehybridised overnight at 70°C in a buffer containing 2X PE (PE = 0.133 M sodium phosphate, 1 mM EDTA), 7% SDS, and 1% BSA. The probe consisted of 25 ng of pUO82 or pUO38 or P2 DNA labelled with α-[^{32}P]dCTP using the Megaprime labelling system (Amersham). The probe was incubated overnight with the membranes at 70°C in the same hybridization buffer. The membranes were then washed for 3 x 15 minutes in 2X SCC, 0.1% SDS at RT and once in 5X SSC, 1% SDS at 70°C and analysed using autoradiography.

Chemiluminenscence probe Southern blot hybridisation with nonradioactive labelling using digoxigenin-11-dUTP was performed with pQE30/OMP26 according to the manufacturer's instructions (DIG-High Prime DNA Labelling and Detection Kit, Boehringer Mannheim). OMP26 is a 26-kDa protein isolated from NTHi-289 that has been investigated as a potential vaccine candidate; DNA sequences encoding OMP26 were previously ligated into pQE30 (El-Adhami *et al.*, 1999). pQE30/OMP26 was kindly provided by Dr Adam Smith, Canberra Centre for Mucosal Immunology.

Chromosomal DNA preparation for PCR Total chromosomal DNA was harvested using the Genomic Tip 20/G kit (Qiagen) from overnight sBHI broth cultures inoculated with a single colony. The DNA concentration was estimated by comparing a small DNA sample with known amounts of bacteriophage lambda DNA on an ethidium bromide stained electrophoretic gel. The DNA was dissolved at a concentration of ~20 ng/μl in TE.

PCR capsular typing A PCR method based on the protocol of Falla *et al.* (1994) was also used to detect the *cap* b locus (Falla *et al.*, 1994). Type b capsule specific primers, B1 and B2, (5'-GCGAAAGTGAACTCTTATCTCTC-3') and (5'-GCTTACGCTTCTATCTC GGTGAA-3'), respectively, were used. The *Taq* PCR Core Kit (Qiagen) was used to prepare the reaction mix to a total volume of 50 μl. It consisted of 1X Qiagen PCR buffer, 3.5 mM $MgCl_2$, 200 μM (each) dNTP mix, 1 μM B1 and 1 μM B2, 2.5 units *Taq* polymerase and 50 ng template DNA. Samples were overlaid with mineral oil and cycled 25 times at 94°C for 1 min, 60°C for 1 min, and 72°C for 1 min, followed by a final 10 min incubation at 72°C. PCR products were run on 1.5% agarose gel containing 0.5 μg/ml ethidium bromide in 1xTAE buffer for 1 hour at 80 V. A negative control, consisting of the same reaction mixture but with no DNA template, was included in each amplification procedure. A standard 1 kb ladder

(Fermentas) was included. The gel images were captured using the Gel Doc 1000 DNA gel analysis and photodocumentation system (Bio-Rad Laboratories, Hercules, CA).

PCR fingerprinting Some of the isolates were PCR fingerprinted using primers directed to enterobacterial repetitive intergenic consensus (ERIC) sequences as described by van Belkum *et al.* (1994) (van Belkum *et al.*, 1994). The *Taq* PCR Core Kit (Qiagen) was used to prepare the reaction mix. The PCR reaction mix (final volume 25 µl) consisted of 2.5 µl 10X Qiagen PCR buffer, 4 µl 25mM $MgCl_2$, 0.5 µl dNTP mix, 1 µl of 25 µM ERIC1 (5'-ATGTAAGCTCCTGGGGATTCAC-3'), 1 µl of 25 µM ERIC2 (5'-AAGTAAGTGACTGGGGTGAGCG-3'), 0.2 µl *Taq* polymerase (1 unit per reaction), 10.8 µl distilled H2O, 5 µl template DNA (~100 ng). PCR samples were overlaid with mineral oil. Samples were cycled 35 times at 94°C for 1 min, 25°C for 1 min, and 74°C for 4 min, followed by a final 10 min incubation at 74°C. PCR products were run on 1.5% agarose gel containing 0.5 µg/ml ethidium bromide in 1xTAE buffer for 3 hours at 50 V. A negative control, consisting of the same reaction mixture but with no DNA template, was included in each amplification procedure. A standard 1 kb ladder (Fermentas) was included. The gel images were captured using the Gel Doc 1000 DNA gel analysis and photodocumentation system (Bio-Rad Laboratories, Hercules, CA). The images were then pasted into Microsoft Powerpoint® for illustration and inclusion in this thesis. Relationship analysis of the fingerprints was performed as described for the PFGE RFLPs using the RAPDistance Programs.

Plasmid analysis Minipreparations of plasmid DNA were prepared by the alkaline lysis method described by Maniatis et al. (Maniatis *et al.*, 1989) and electrophoresed in 1% agarose in 1x TAE electrophoresis buffer under standard electrophoresis conditions. Purified plasmid, pQE30/OMP26, was prepared using the Quantum Prep® Plasmid Miniprep Kit (Bio-Rad) by the protocol supplied by the manufacturer.

Beta-lactasmase testing Beta-lactamase production was detected with nitrocefin-impregnated paper discs (Cefinase discs; BBL).

Diplomo Diplomo is an interactive computer graphics application that uses the distance plot principle in an approach to phylogenetics analysis. Diplomo compares different distance measures with each other by displaying them as a scatter plot. It then helps one identify individual comparisons within the plot. Many different questions can be

addressed with a DIPLOMO analysis and it was used here to determine whether distance matrices derived from *Sma*I digestion of Hib DNAs varied from distance matrices derived from *Apa*I digestion of Hib DNAs. Diplomo is distributed by the Research School of Biological Sciences, Australian National University (Weiller and Gibbs, 1993).

Simpson's index of diversity Simpson's index of diversity (D) is a numerical index of discrimination based on the probability that two unrelated strains sampled from the test population will be placed into different typing groups. This probability is calculated as

$$D = 1 - \frac{1}{N(N-1)} \sum_{j=1}^{S} n_j (n_j - 1)$$

where N is the total number of strains in the sample population, S is the total number of types described, and n_j is the number of strains belonging to the jth type (Hunter and Gaston, 1988). The application of this index was used to assess the discriminatory power of *Sma*I and *Apa*I PFGE.

GenStat (Release 6.1) GenStat is a comprehensive statistical system developed by Rothamsted Experimental Station and distributed by the Numerical Algorithms Group, United Kingdom. In appropriate cases, average distances or chi square analysis was used. Random average distances between types associated with either meningitis or epiglottitis were determined to evaluate the association between type and disease manifestation. Pearson chi square values at $p<.001$ were calculated to evaluate the association between epidemiologic factors and RFLP type. Consultation and the use of GenStat was made available through the Statistical Consulting Unit, The Graduate School, Australian National University.

Methods Addendum

Bacterial isolates. Fifty Hib isolates collected after the introduction of vaccines in northern Australia were obtained through the courtesy of Heidi C. Smith-Vaughan at the Menzies School of Health Research. Twenty of these isolates have been PFGE genotyped and the results are included in Chapter 8. The gel images were captured using the Gel Doc 1000 DNA gel analysis and photodocumentation system (Bio-Rad Laboratories, Hercules, CA).

CHAPTER 3

Evaluation of Pulsed Field Gel Electrophoresis to Investigate the Genetic Diversity of *Haemophilus influenzae* type b

3.1 Introduction

Two of the most common manifestations of Hib disease are meningitis and epiglottitis in children. As described in Chapter 1 the following intriguing issues surround the epidemiology of these diseases:

1) In Western communities and indigenous populations these infections tend to occur at different ages according to the risk of Hib disease in the population. The absence of epiglottitis among Aborigines corresponds to this striking characteristic of age-related variation in the incidence of Hib disease (Clements and Gilbert, 1989). It is not known why an age related incidence of disease occurs.

2) The ability of Hib to infect the epiglottis is characteristic as epiglottitis in children is rarely attributed to other microorganisms (Hanna *et al.*, 1992). It is not known why this is so or if specific strains of Hib are associated with epiglottitis.

3) A pattern of Hib disease in Australia before the introduction of Hib vaccines was a significant difference in the incidence of epiglottitis among urban populations. The incidence of epiglottitis in Victoria (Gilbert *et al.*, 1990) and the Australian Capital Territory (McGregor *et al.*, 1992) was twice that in Sydney (McIntyre *et al.*, 1991) and Western Australia (Hanna *et al.*, 1992). Whether differences in the incidence of epiglottitis in different populations are related to different strains of Hib is not known.

4) It is very rare for a child to develop concurrent Hib infections of the epiglottis and the meninges (personal communication, G. Gilbert). The reasons for this are unknown nor is it known why children develop epiglottitis instead of meningitis or why contacts of children with meningitis may develop epiglottitis. While it has been reported that some Hib outer membrane protein (OMP) subtypes are more likely to cause meningitis than epiglottitis (Takala *et al.*, 1987) this finding has not been substantiated by later studies (Takala *et al.*, 1989; Clements *et al.*, 1992). To date no finding has defined the characteristics of Hib, if any, that are associated with its unique epidemiology.

In order to investigate the possibility that bacterial characteristics may account for its unique epidemiology, I designed a preliminary study to examine the genotypic structure of Hib isolates in a sample of meningitis and epiglottitis isolates recovered from non-

Aboriginal patients admitted to the Canberra Hospital during the period 1988-1990. When this study commenced in 1992, pulsed field gel electrophoresis was not well established and had not previously been used to characterise a large number of Hib isolates. It was necessary to establish the optimal conditions for use of PFGE and evaluate the methodology for use in examining Hib isolates. This chapter describes the results of the first attempt to study Hib isolates using PFGE in order to answer questions concerning its epidemiology and, specifically, to answer questions relating to isolate types associated with epiglottitis and meningitis.

The following hypothesis concerning the methodology was tested: that PFGE of Hib genomic DNAs digested with the restriction endonucleases, *Sma*I or *Apa*I, would produce a limited number of well resolved fragments that could be used to reliably discriminate between closely related and more distantly related isolates of Hib.

Two hypotheses concerning the bacterial isolates were tested: first, that genotypes associated with epiglottitis are not associated with meningitis and, secondly that genotypes associated with meningitis are not associated with epiglottitis.

3.2 Methods

This experiment was performed after an extended period spent developing the protocol and my proficiency in using PFGE. The particulars of the methods are described below and in more detail in Chapter 2.

3.2.1 Bacterial isolates (Appendix A.1)

A selection of 20 isolates recovered from patients with meningitis and 14 isolates recovered from patients with epiglottitis from the Canberra region were examined in a preliminary evaluation of the use of PFGE for characterising Hib. The 34 isolates were collected between November 1988 and September 1990 as part of a national collaborative study of the antimicrobial resistance in Australia among *Haemophilus influenzae* isolates from blood or cerebrospinal fluid (Collignon *et al.*, 1992). All of the isolates had previously been identified by conventional methods.

The meningitis isolates were designated M1 to M20 and the epiglottitis isolates were designated E1 to E15 excluding E8, an isolate that was subsequently found to have been recovered from a patient from Campbelltown, NSW.

3.2.2 Restriction endonuclease digestion

Initially, *Sma*I (CCC↓↑GGG) and *Apa*I (G↑GGCC↓C) were used. Washing of plugs and digestion of the DNA with endonucleases were performed as described in Methods.

3.2.3 PFGE conditions

Separation of fragments using a clamped homogeneous electric field device was carried out as described in Methods.

3.2.4 Estimation of genetic diversity between isolates and construction of dendrograms

The numbers and mobilities of fragments were determined by visual examination of the polaroid photographs of the stained gels as described in Methods. A matrix of F values for all pairs of isolates was constructed and a dendrogram computed using the NJTREE program and the companion TDRAW program.

3.2.5 Hybridisation with pU082 and P2 probes

Hybridisation techniques were used to detect the presence of the *cap* locus and the P2 gene using the *Sma*I PFGE gels as described in Chapter 2.

3.2.6 Correlation of data and determination of the discriminatory power of enzymes

Visual analysis and computer analysis using the programs Diplomo (Weiller and Gibbs, 1993) and the RAPDistance Programs were used to evaluate the correlation of the two enzymes. Simpson's Index of Diversity (Hunter and Gaston, 1988) was used to determine their discriminatory power.

3.2.7 Average distance calculations of RFLP type and disease association

The distance matrix of the 26 SmaI-RFLPs types found among the 34 meningitis and epiglottitis isolates (Table 3.9 in Appendix B) was tested using GenStat's average distances program for evidence of association of type to disease manifestation.

3.3 Results

The results are presented in layers of information to illustrate the evaluation of the effect of sample size and selection and the effect of using the different enzymes. First, the analysis of the RFLP patterns by visual comparison is described for these subsets: the *Sma*I-RFLP patterns, the *Apa*I-RFLP patterns, and the RFLP patterns found when the *Sma*I fragments and the *Apa*I fragments of each isolate are combined. Each of these subsets is examined for the sample of meningitis isolates, the epiglottitis isolates, and the combined sample of meningitis and epiglottitis isolates. Secondly, the genetic diversity determined by numerical analysis of the fragment patterns and the construction of dendrograms is similarly described for each of the subsets. Table 3.5 summarises the RFLP types found for each enzyme and diagnosis.

RFLP types are designated by the restriction enzyme(s) and the lowest M then E numbered isolate found with that pattern. For example, *Sma*I-M5 includes isolates M5, M12, M16, M18, and E12.

3.3.1 Hybridisation

All 34 isolates hybridised with pU082. Hybridisation was detected on one or two DNA fragments per isolate on the *Sma*I-PFGE gels. The same size fragment, corresponding to fragment 8.1 (~416 kb) of HS008, of 21 types hybridised with pU082. Hybridisation was detected on two fragments of 13 types. Among sets of multiple types the same fragment or fragments hybridised and no further splitting of RFLP types was found from the capsular typing of the PFGE RFLPs.

All the isolates hybridised with the P2 probe. Hybridisation was found on one fragment per isolate and for 21 types the same size fragment, corresponding to fragment 8.6 (~107 kb) of HS008, hybridised. A fragment corresponding to fragment 8.14 (~35 kb) of HS008 hybridised with four clonally related types, M1, M3, M20 and E15. Hybridisation also occurred on an ~70 kb fragment of M13, the most diverse isolate in the sample. No further splitting of RFLP types was found using the P2 probe.

3.3.2 Visual analysis of *Sma*I and *Apa*I RFLPs

Well resolved patterns were found among the isolates with both enzymes but the fragments of *Sma*I digested DNAs were more distinct. All isolates were digested at least twice and usually three or more times in the development of the methodology to

confirm fragment size. The size range of the fragments found among the *Sma*I-RFLP types was ~9–420 kb and ~6-450 kb among the *Apa*I-RFLP types. Two *Sma*I fragments were shared by all the isolates. Neither of those shared fragments hybridised with pU082 or the P2 gene probe. Seven of the *Apa*I fragments were found among all the isolates. Comparative data on the fragments is shown in Table 3.1. Tables 3.2-3.4 describe the fragment distribution among types found in each subset. The RFLP patterns were stable. When isolates were repeatedly subcultured, their DNAs yielded unchanged *Sma*I and *Apa*I fragment patterns. Isolate HS008 was most often subcultured and repeatedly digested and no variation was seen in its RFLP pattern.

Table 3. 1 Comparative data on the PFGE fragments from *Sma*I and *Apa*I digestion of the DNAs of 20 meningitis and 14 epiglottitis isolates of *Haemophilus influenzae* type b

	Meningitis			Epiglottitis			Meningitis and Epiglottitis		
	*Sma*I	*Apa*I	*Sma*I/ *Apa*I	*Sma*I	*Apa*I	*Sma*I/ *Apa*I	*Sma*I	*Apa*I	*Sma*I/ *Apa*I
Number of types	16	15	17	12	9	13	26	20	28
Number of distinct fragments[a]	48	42	90	36	31	69	49	44	97
Total number of fragments[b]	232	275	555	173	159	417	377	363	909
Average number of fragments per type	14.5	18.3	32.6	14.4	17.7	32.0	14.5	18.2	32.5

[a,b]Refers to fragments among the different types found in the subset (multiple isolates of the same type are not included)

Meningitis Isolates: *Sma*I-RFLPs (Figures 3.1-3.2, Table 3.2) Sixteen *Sma*I-RFLP types were found among the 20 meningitis isolates. One type, *Sma*I-M5, comprising 4 of the isolates, accounted for 20% of the sample. The distribution of the meningitis isolates among types was:

*Sma*I-RFLP type	Number of isolates/type
M5	4
M3	2
all others *(n=14)*	1

Epiglottitis Isolates: *Sma*I-RFLPs (Figures 3.2-3.3, Table 3.2) Twelve *Sma*I-RFLP types were found among the 14 epiglottitis isolates. Two types, *Sma*I-E2 and *Sma*I-E9, each comprised 2 (14%) of the isolates. The distribution of epiglottitis isolates among *Sma*I-RFLP types was:

*Sma*I-RFLP type	Number of isolates/type
E2	2
E9	2
all others (*n=10)*	1

Meningitis and Epiglottitis Isolates Combined: *Sma*I-RFLPs (Table 3.2) Twenty-six *Sma*I-RFLP types were found in the combined sample. Only 2 *Sma*I-RFLP types were found in both disease groups. One of these, *Sma*I-M5, accounted for 15% of the sample. The distribution of the combined sample of meningitis and epiglottitis isolates among the *Sma*I-RFLP types was:

*Sma*I-RFLP type	Number of isolates/type	Distribution of isolates/type/disease	
		Meningitis	Epiglottitis
M5	5	4	1
M8	2	1	1
M3	2	2	—
E2	2	—	2
E9	2	—	2
all others (*n=21)*	1	—	—

Meningitis Isolates: *Apa*I-RFLPs (Figures 3.4-3.5, Table 3.3) Fifteen *Apa*I-RFLP types were found among the 20 meningitis isolates. The RFLP type designated *Apa*I-M5 accounted for 30% of the sample. The meningitis isolates were distributed among the *Apa*I-RFLP types as follows:

*Apa*I-RFLP type	Number of isolates/type
M5	6
all others *(n=14)*	1

Epiglottitis Isolates: *Apa*I-RFLPs (Figures 3.5-3.6, Table 3.3) Nine *Apa*I-RFLP types were found among the 14 epiglottitis isolates. RFLP type *Apa*I-E6 accounted for 31% of the sample. The isolates were distributed among the *Apa*I-RFLP types as follows:

*Apa*I-RFLP type	Number of isolates/type
E6	5
E5	2
all others (*n=7)*	1

Meningitis and Epiglottitis Isolates Combined: *Apa*I-RFLPs (Table 3.3) Twenty *Apa*I-RFLP types were found in the combined sample of meningitis and epiglottitis isolates. Four types were found in both disease groups. A predominant type, *Apa*I-M5, comprised 32% of the sample. The combined meningitis and epiglottitis isolates were distributed among *Apa*I-RFLP types as follows:

*Apa*I-RFLP type	Number of isolates per type	Distribution of isolates/type/disease	
		Meningitis	Epiglottitis
M5	11	6	5
M15	3	1	2
M8	2	1	1
M20	2	1	1
all others *(n=16)*	1	—	—

Meningitis Isolates: *Sma*I/*Apa*I-RFLPs (Table 3.4) The *Sma*I and *Apa*I fragments were combined for each of the meningitis isolates. Seventeen *Sma*I/*Apa*I-RFLP types were found among the 20 meningitis isolates. *Sma*I/*Apa*I-M5 accounted for 20% of the sample. The distribution of the isolates among the types was:

*Sma*I/*Apa*I-RFLP type	Number of isolates/type
M5	4
all others (*n=16)*	1

Epiglottitis Isolates: *Sma*I/*Apa*I-RFLPs (Table 3.4) The *Sma*I and *Apa*I fragment were combined for each of the epiglottitis isolates. Thirteen *Sma*I/*Apa*I-RFLP types were found among the 14 epiglottitis isolates. Only one type, *Sma*I/*Apa*I-E9, with 2 members represented multiple isolates. The distribution of isolates among types was:

*Sma*I/*Apa*I-RFLP type	Number of isolates/type
E9	2
all others (*n=12)*	1

Combined Meningitis and Epiglottitis Isolates: *Sma*I/*Apa*I-RFLPs (Table 3.4) Twenty-eight *Sma*I/*Apa*I-RFLP types were found. *Sma*I/*Apa*I-M5 accounted for 15% of

the sample. The combined meningitis and epiglottitis isolates were distributed among the *Sma*I/*Apa*I-RFLP types as follows:

*Sma*I/*Apa*I RFLP type	Number of isolates/type	Distribution of isolates/type/disease	
		Meningitis	Epiglottitis
M5	5	4	1
M8	2	1	1
E9	2	—	2
all others *(n=25)*	1	—	—

Table 3. 2 Number of *Sma*I-RFLP types and number of fragments per type

	Total number of types		Number of fragments per type			
			13	14	15	16
Meningitis isolates *n=20*	16 (0.80:1)[a]	Types	2	6	6	2
		Isolates	*n=5*	*n=7*	*n=6*	*n=2*
Epiglottitis isolates *n=14*	12 (0.86:1)[a]	Types	2	5	3	2
		Isolates	*n=3*	*n=6*	*n=3*	*n=2*
All isolates *n=34*	26[b] (0.76:1)[a]	Types	3[c]	11	8[d]	4
		Isolates	*n=8*	*n=13*	*n=9*	*n=4*

[a]Ratio of the number of types found to the total number of isolates typed.
[b]Two *Sma*I-RFLP types were found among both meningitis and epiglottitis isolates. Thus, the total number of different *Sma*I-RFLP types found in the sample was 26 not 28.
[c]One of the 13 fragment *Sma*I-RFLPs was found among both groups of isolates.
[d]One of the 15 fragment *Sma*I-RFLPs was found among both groups of isolates.

Table 3. 3 Number of *Apa*I-RFLP types and number of fragments per type

	Total number of types		Number of fragments per type					
			16	17	18	19	20	21
Meningitis isolates *n=20*	15 (0.75:1)[a]	Types		5	3	5	1	1
		Isolates		*n=10*	*n=3*	*n=5*	*n=1*	*n=1*
Epiglottitis isolates *n=14*	9 (0.64:1)[a]	Types	1	3	3	2		
		Isolates	*n=1*	*n=6*	*n=4*	*n=3*		
All isolates *n=34*	20[b] (0.71:1)[a]	Types	1	6[c]	5[d]	6[e]	1	1
		Isolates	*n=1*	*n=16*	*n=7*	*n=8*	*n=1*	*n=1*

[a]Ratio of the number of types found to the total number of isolates typed.
[b]Four *Apa*I-RFLP types were found among both meningitis and epiglottitis isolates.
[c]Two of the17 fragment *Apa*I-RFLPs were found among both groups of isolates.
[d]One of the 18 fragment *Apa*I-RFLPs was found among both groups of isolates.
[e]One of the 19 fragment *Apa*I-RFLPs was found among both groups of isolates.

Table 3. 4 Number of *Sma*I/*Apa*I-RFLP types and number of fragments per type

	Total number of types		Number of fragments per type						
			30	31	32	33	34	35	37
Meningitis isolates *n=20*	17 (0.90:1)[a]	Types	2	1	6	4	2	1	1
		Isolates	*n=5*	*n=1*	*n=6*	*n=4*	*n=2*	*n=1*	*n=1*
Epiglottitis isolates *n=14*	13 (0.93:1)[a]	Types	2	1	5	4	1		
		Isolates	*n=3*	*n=1*	*n=5*	*n=4*	*n=1*		
All isolates *n=34*	28[b] (0.71:1)[a]	Types	3[c]	2	10[d]	8	3	1	1
		Isolates	*n=8*	*n=2*	*n=11*	*n=8*	*n=3*	*n=1*	*n=1*

[a]Ratio of the number of types found to the total number of isolates typed.
[b]Two RFLP types were found among both meningitis and epiglottitis isolates. Thus, the total number of different *Sma*I/*Apa*I RFLP types found in the sample was 28 not 30.
[c]One of the 30 fragment RFLPs was found among both the meningitis and epiglottitis isolates.
[d]One of the 32 fragment RFLPs was found among both the meningitis and epiglottitis isolates.

3.3.3 Genetic relationships determined by numerical analysis of fragment length polymorphisms

As described in the Methods section in Chapter 2, in order to quantify the genetic relationships among the RFLP types similarity matrices were built using Nei and Li's *F* values obtained by comparing pairs of types. Figures 3.7 to 3.15 are the dendrograms generated from the matrices built in this study.

The distance matrices are included in Appendix B to illustrate the data from which dendrograms are computed. When matrices are built for a limited number of isolates it is easy to pick out the most closely related and the most distantly related isolates from the data. However, as matrices enlarge owing to increasing numbers of isolates it becomes increasingly difficult to utilise them to visualise the genetic distance between pairs. The construction of dendrograms helps to overcome this problem. Table 3.6 summarises numerical data derived using Nei and Li's mathematical model.

Meningitis Isolates: *Sma*I-RFLPs (Figure 3.7, Table 3.7 in Appendix B) Table 3.7 shows the distance matrix of 1-*F* values derived from the pairwise comparison of the 16 types found among the 20 meningitis isolates and from which the dendrogram seen in Figure 3.7 was constructed. The similarity among pairs ranged from 19% to 97%. RFLP type *Sma*I-M5 comprising 4 of the isolates was clonally related to 6 other types (*Sma*I-M2, M3, M8, M11, M17, and M19) that comprised 7 isolates.

...text continues after Figures 3.1 – 3.6

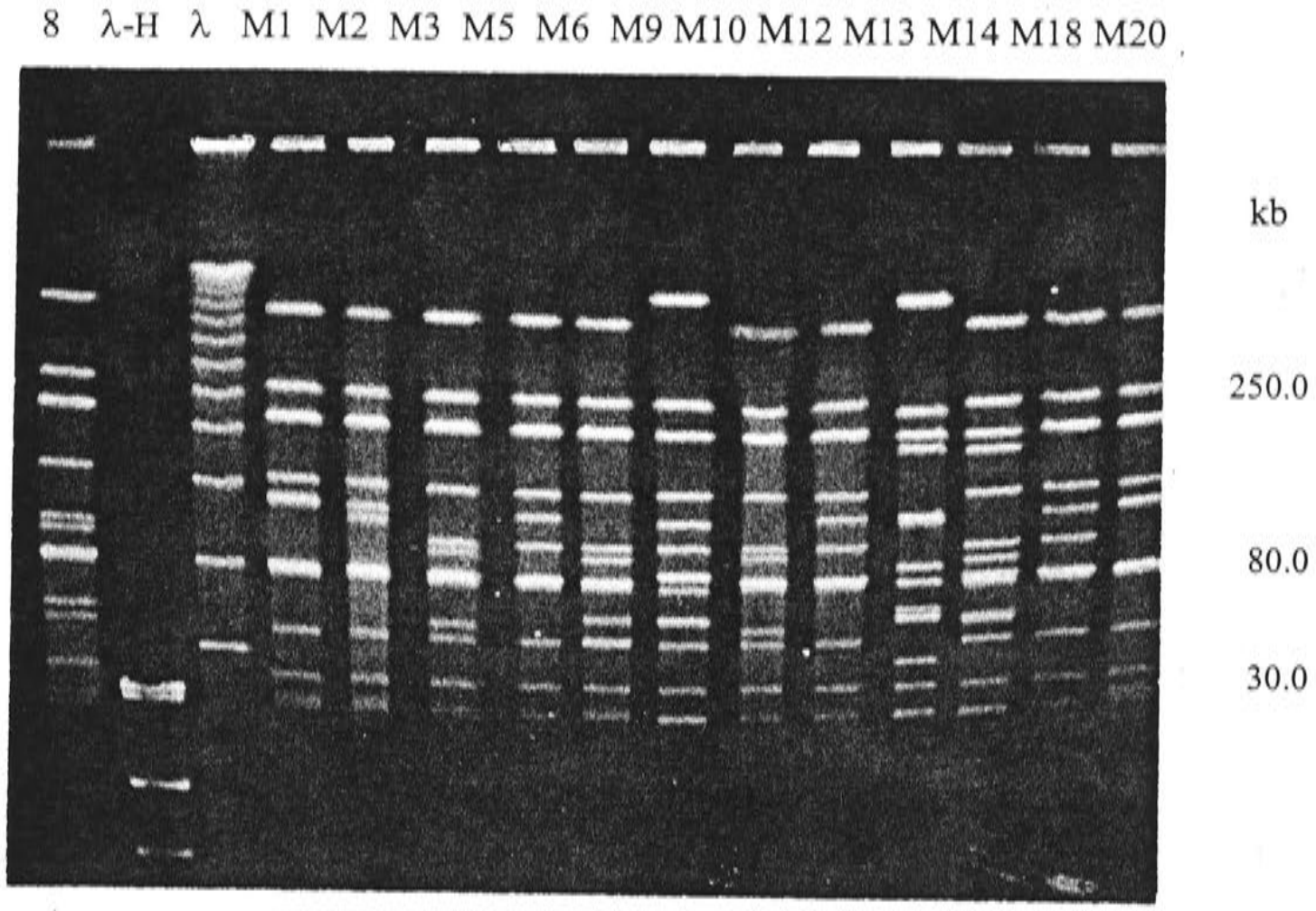


Figure 3.1 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 12 Hib isolates recovered from the cerebrospinal fluid or blood of meningitis patients from the Canberra region. Lanes labelled 8, λ, and λ-H contain fragment standards.

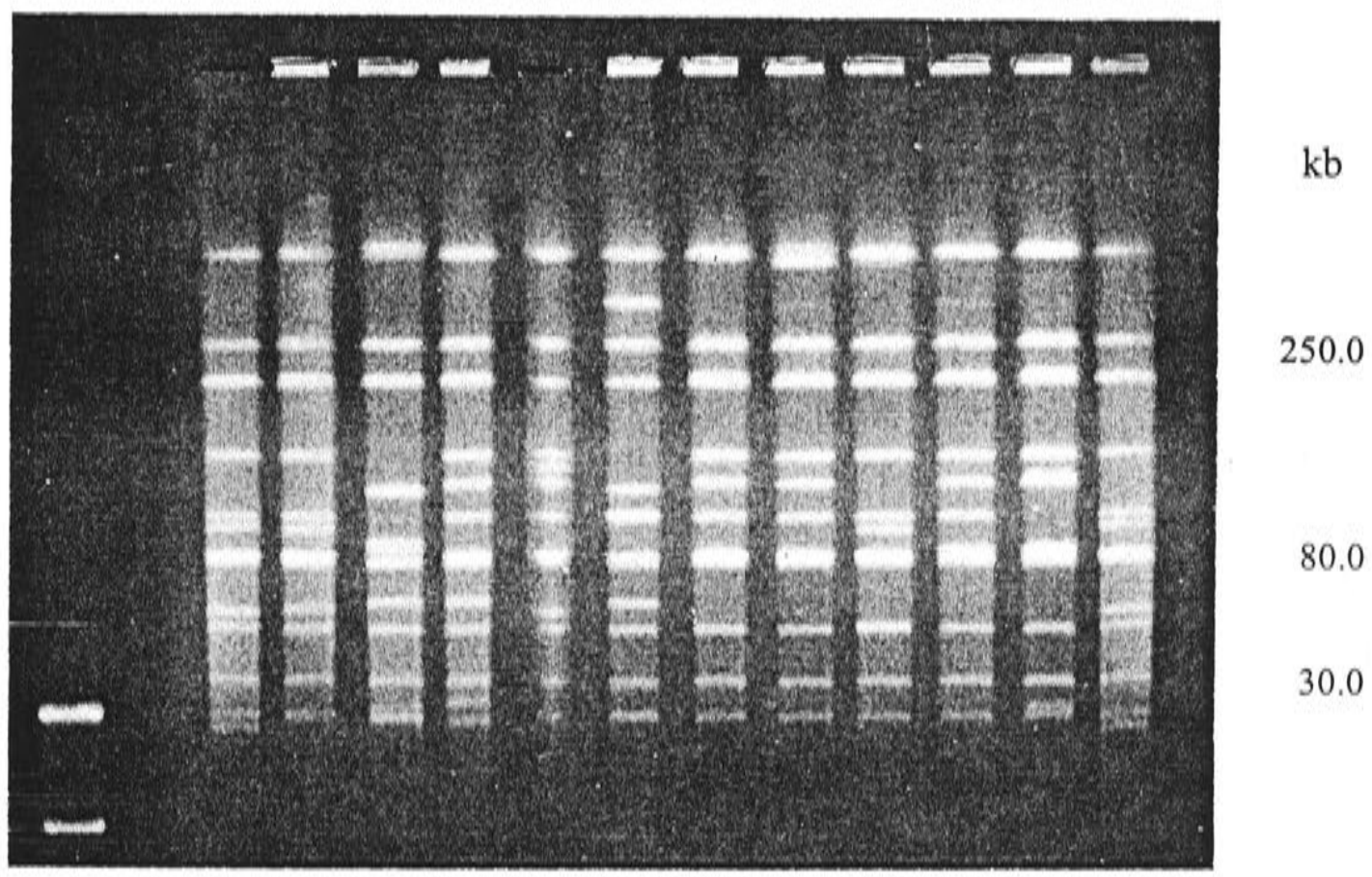


Figure 3.2 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 10 Hib isolates recovered from the cerebrospinal fluid or blood of 8 meningitis patients and 2 epiglottitis patients from the Canberra region. Lanes labelled 8 and λ-H contain fragment standards.

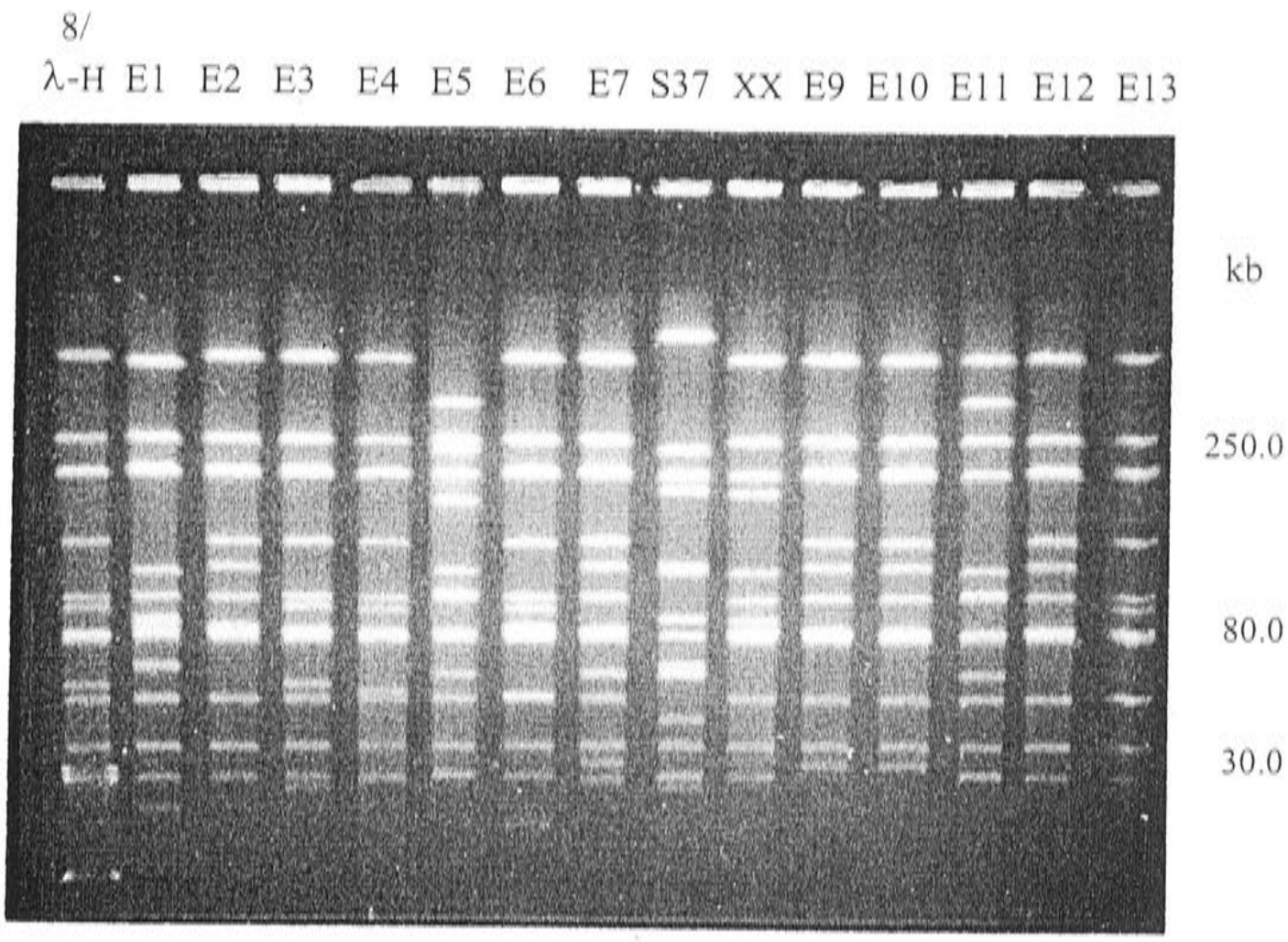


Figure 3.3 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 13 Hib isolates recovered from the blood of epiglottitis patients from the Canberra region and 1 isolate from the Sydney region. The lane labelled 8/λ-H contains fragment standards. The lane labelled XX is an isolate that is not part of this study.

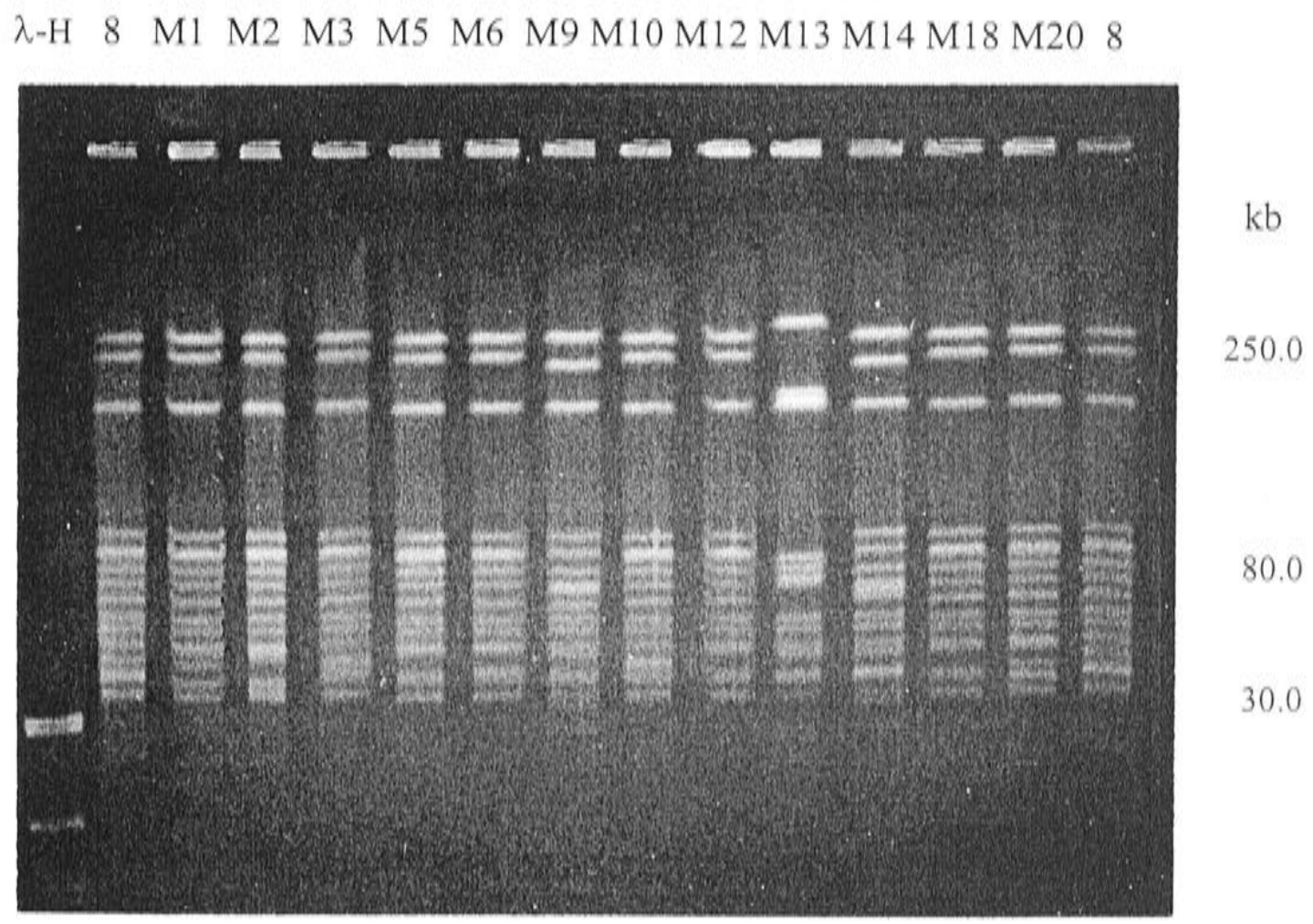


Figure 3.4 *Apa*I restriction fragments obtained by PFGE of the genomic DNAs of 12 Hib isolates recovered from the cerebrospinal fluid or blood of meningitis patients from the Canberra region. Lanes labelled 8 and λ-H contain fragment standards.

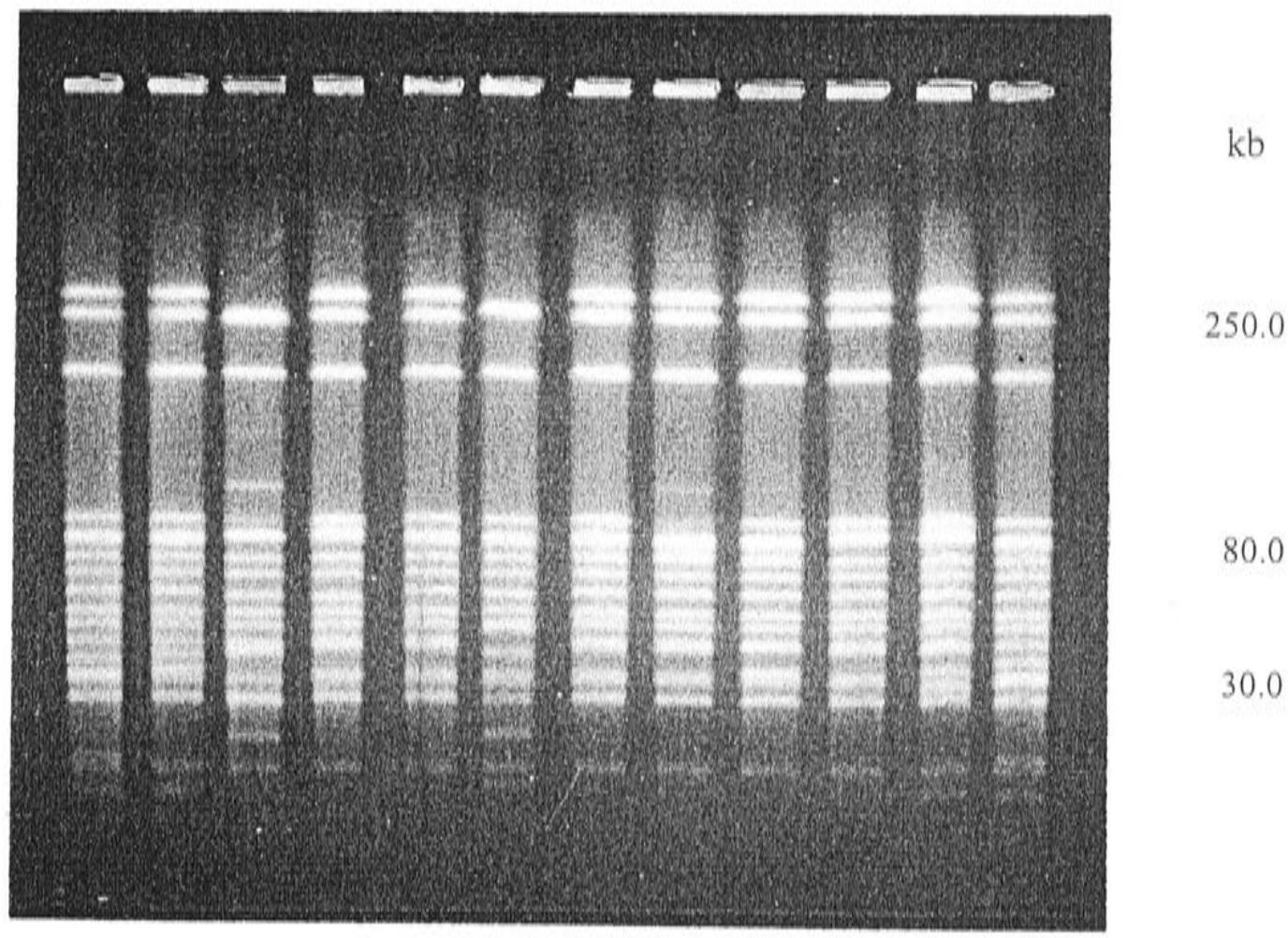


Figure 3.5 *Apa*I restriction fragments obtained by PFGE of the genomic DNAs of 12 Hib isolates recovered from the cerebrospinal fluid or blood of 8 meningitis and 2 epiglottitis patients from the Canberra region. Lanes labelled 8 contain fragment standards.

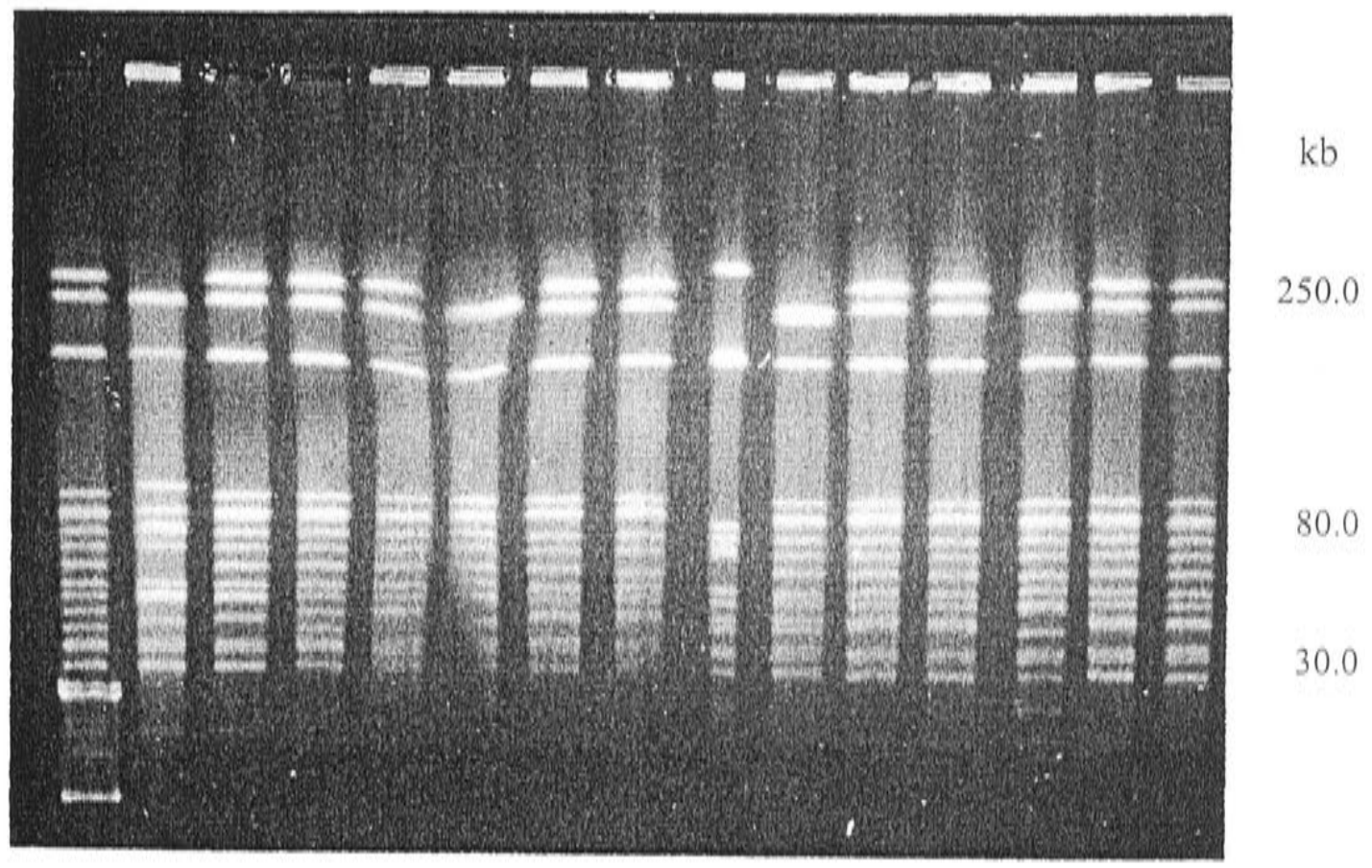


Figure 3.6 *Apa*I restriction fragments obtained by PFGE of the genomic DNAs of 12 Hib isolates recovered from the blood of epiglottitis patients from the Canberra region and 1 epiglottitis isolate from the Sydney region. The lane labelled 8/λ-H contains fragment standards. The lane labelled XX is an isolate that is not part of this study.

Thus, *Sma*I-M5 and its clonally related types accounted for 55% of the sample. *Sma*I-M7, *Sma*I-M9, and *Sma*I-M13 defined the range of diversity within the sample relative to the predominant type. Each represented a single isolate that was less than 43% (range 36-43%) similar to *Sma*I-M5.

Epiglottitis Isolates: *Sma*I-RFLPs (Figure 3.8, Table 3.8 in Appendix B) Table 3.8 shows the distance matrix of 1-*F* values derived from the pairwise comparison of the 12 *Sma*I-RFLP types found among the 14 epiglottitis isolates and from which the dendrogram seen in Figure 3.8 was constructed. The similarity among pairs ranged from 48% to 96%. One of the clonal groups included 4 types (*Sma*I-E2, *Sma*I-E7, *Sma*I-E9, and *Sma*I-E12) and comprised 6 members (43%) of the sample. *Sma*I-E5 was the most genetically diverse type with *F* values ranging from 0.51 to 0.62 between it and the predominant clonal types.

Meningitis and Epiglottitis Isolates Combined: *Sma*I-RFLPs (Figure 3.9, Table 3.9 in Appendix B) Table 3.9 shows the distance matrix of 1-*F* values derived from the pairwise comparison of the 26 *Sma*I-RFLP types found among the 34 meningitis and epiglottitis isolates and from which the dendrogram seen in Figure 3.9 was constructed. The similarity among pairs ranged from 19% to 97%. Six of the clusters at an *F* value ≥ 0.9 contained both meningitis and epiglottitis isolates; however only two RFLP types, *Sma*I-M5 and *Sma*I-M8, comprised both meningitis and epiglottitis isolates.

Meningitis Isolates: *Apa*I-RFLPs (Figure 3.10, Table 3.10 in Appendix B) Table 3.10 shows the distance matrix of 1-*F* values derived from the pairwise comparison of the 12 types found among the 20 meningitis isolates and from which the dendrogram seen in Figure 3.10 was constructed. The similarity among pairs ranged from 43% to 97%. RFLP type *Apa*I-M5 and 3 clonally related types accounted for 45% of the sample. *Apa*I-M13 representing a single isolate was the most genetically diverse type from the *Apa*I-M5 type at an *F* value of 0.56.

Epiglottitis Isolates: *Apa*I-RFLPs (Figure 3.11, Table 3.11 in Appendix B) Table 3.11 shows the distance matrix of 1-*F* values derived from the pairwise comparison of the 9 types found among the 14 epiglottitis isolates and from which the dendrogram seen in Figure 3.11 was constructed. The similarity among pairs ranged from 53% to 97%. RFLP type *Apa*I-E6 and 3 clonally related types accounted for 57% of the sample.

Table 3. 5 Summary of RFLP types* found per enzyme and diagnosis

	Meningitis RFLPs *n=20*			Epiglottitis RFLPs *n=14*			Meningitis and Epiglottitis Combined RFLPs *n=34*		
	*Sma*I	*Apa*I	*Sma*I/ *Apa*I	*Sma*I	*Apa*I	*Sma*I/ *Apa*I	*Sma*I	*Apa*I	*Sma*I/ *Apa*I
1	M1	M1	M1	E1	E1	E1	M1	M1	M1
2	M2	M2	M2	**E2**[d]	E2	E2	M2	M2	M2
3	**M3**[a]	M3	M3	E3	E3	E3	**M3**[a]	M3	M3
4	**M5**[b]	M4	M4	E4	E4	E4	**M5**[h]	M4	M4
5	M6	**M5**[c]	**M5**[b]	E5	**E5**[f]	E5	M6	**M5**[j]	**M5**[h]
6	M7	M7	M6	E6	**E6**[g]	E6	M7	M7	M6
7	M8	M8	M7	E7	E7	E7	**M8**[i]	**M8**[i]	M7
8	M9	M9	M8	**E9**[e]	E13	**E9**[e]	M9	M9	**M8**[i]
9	M10	M10	M9	E11	E15	E11	M10	M10	M9
10	M11	M13	M10	E12		E12	M11	M13	M10
11	M13	M14	M11	E13		E13	M13	M14	M11
12	M14	M15	M13	E15		E14	M14	**M15**[k]	M13
13	M15	M17	M14			E15	M15	M17	M14
14	M17	M19	M15				M17	M19	M15
15	M19	M20	M17				M19	**M20**[l]	M17
16	M20		M19				M20	E1	M19
17			M20				E1	E2	M20
18							**E2**[d]	E3	E1
19							E3	E4	E2
20							E4	E13	E3
21							E5		E4
22							E6		E5
23							**E9**[e]		E6
24							E11		**E9**[e]
25							E13		E11
26							E15		E13
27									E14
28									E15

[a]M3 and M4
[b]M5,M12,M16,M18
[c]M5,M6,M11,M12,M16, M18
[d]E2 and E14
[e]E9 and E10
[f]E5 and E11
[g]E6,E9,E10,E12,E14
[h]M5,M12,M16,M18,E12
[i]M8 and E7
[j]M5,M6,M11,M12,M16, M18,E6,E9,E10,E12, E14
[k]M15,E5, E11
[l]M20 and E15

*The designated RFLP type name represents the lowest number assigned to a meningitis isolate of that type or if no meningitis isolate represents the type then the lowest number assigned to an epiglottitis isolate of that type. The types in boldface represent multiple isolates. The boxed types represent types found among both the meningitis and the epiglottitis isolates. Note: In the text of this thesis *Sma*I, *Apa*I or *Sma*I/*Apa*I prefixes RFLP types as appropriate for clarification.

Meningitis and Epiglottitis Isolates Combined: *Apa*I-RFLPs (Figure 3.12, Table 3.12 in Appendix B) Table 3.12 shows the distance matrix of 1-*F* values derived from the pairwise comparison of the 20 types found among the 34 meningitis and epiglottitis isolates and from which the dendrogram seen in Figure 3.12 was constructed. The similarity among pairs ranged from 43% to 97%. Three of the clusters at an *F* value ≥ 0.9 and four types comprised both meningitis and epiglottitis isolates. *Apa*I-M5 and clonally related types accounted for 50% of the sample. *Apa*I-M13 was the most diverse isolate from these types.

Meningitis Isolates: *Sma*I/*Apa*I-RFLPs (Figure 3.13, Table 3.13 in Appendix B) Table 3.13 shows the distance matrix of 1-*F* values derived from the pairwise comparison of the 17 *Sma*I/*Apa*I-RFLP types found among the 20 meningitis isolates and from which the dendrogram seen in Figure 3.13 was constructed. The similarity among pairs ranged from 36% to 97%. Only one type, *Sma*I/*Apa*I-M5, with 4 members represented multiple isolates.

Epiglottitis Isolates: *Sma*I/*Apa*I-RFLPs (Figure 3.14, Table 3.14 in Appendix B) Table 3.14 shows the distance matrix of 1-*F* values derived from the pairwise comparison of the 13 *Sma*I/*Apa*I-RFLP types found among the 14 epiglottitis isolates and from which the dendrogram seen in Figure 3.14 was constructed. The similarity among pairs ranged from 55% to 97%. Only one type, *Sma*I/*Apa*I-E9 that had two members, represented more than one isolate.

Combined Meningitis and Epiglottitis Isolates: *Sma*I/*Apa*I-RFLPS (Figure 3.15, Table 3.15 in Appendix B) Table 3.15 shows the distance matrix of 1-*F* values derived from the pairwise comparison of the 28 *Sma*I/*Apa*I-RFLP types found among the combined meningitis and epiglottitis isolates and from which the dendrogram seen in Figure 3.15 was constructed. The similarity among pairs ranged from 36% to 98%. Only 2 types comprised both meningitis and epiglottitis isolates, however, 6 of the clusters at an *F* value ≥ 0.9 contained both meningitis and epiglottitis isolates. *Sma*I/*Apa*I-M5 was indistinguishable from 4 other isolates and found among both the meningitis and epiglottis isolates. It was closely or clonally related (F≥0.90) to 11 isolates in three clonal groups. Thus, M5 and isolates clonally or closely related accounted for 47% of the sample.

3.3.4 Genetic diversity and correlation to meningitis and epiglottitis

Clustering of the RFLP types was not associated with either meningitis or epiglottitis. Isolates from patients with either disease type were represented by a diverse range of types and no single type was associated with a particular diagnosis. Chi square analysis did not reveal any association. In addition, no obvious association was found between age, sex, or date of collection.

3.3.5 Correlation of *Sma*I and *Apa*I RFLP types

Visual examination of the matrices, dendrograms, and the patterns of the RFLP types showed that overall relationships between isolates were similar using both enzymes and that both meningitis and epiglottitis isolates shared RFLP types. The correlation coefficient of the two enzymes determined using the computer program Diplomo (Weiller and Gibbs, 1993) was 0.79 among the meningitis isolates (190 data points), 0.74 among the epiglottitis isolates (91 data points), and 0.78 among the combined sample (561 data points).

Specific differences in *Sma*I versus *Apa*I typing could be reliably ascertained only among the types with multiple isolates. In almost all cases typing differences were due to indistinguishable isolates being identified as clonal or closely related by the second enzyme. Examples of these typing differences are described in the box following and are illustrated in the summary of types shown in Table 3.5.

3.3.6 Numerical index of the discriminatory ability of the typing system

Simpson's index of diversity (D) was applied to the data to determine the discriminatory power of *Sma*I and *Apa*I PFGE. The use of *Sma*I alone had an index of 0.974. This index indicates that if two strains were sampled randomly from the population, than on 97.4% of occasions, they would fall into different types. The index for *Apa*I alone was 0.890, a D value considered to be less desirable. Using the two enzymes together yielded an index of 0.978, marginally better than *Sma*I alone. No correcting factor for the small sample size ($n=34$) was applied in determining D.

3.3.7 Average distances calculation of RFLP type and disease association

The distance matrix of the 26 *Sma*I-RFLPs (Table 3.9 in Appendix B) was tested for evidence of association of type to disease manifestation by determining average

distances between types. The average distance between meningitis isolates was 0.3475, between epiglottitis isolates it was 0.3000, and between meningitis and epiglottis isolates the average distance was 0.3160. Permutation test estimates from 199 permuted samples were 0.3245, 0.3275, and 0.3258, respectively. The results were well within the 95% confidence level and, thus, no evidence was provided to suggest that any types were associated with disease manifestation.

***Sma*I vs *Apa*I typing:** Differences among types with **multiple** isolates

Meningitis

- M3 and M4, indistinguishable by *Sma*I typing, were different though closely related by *Apa*I typing.
- *Apa*I-M5, included all the isolates that fell in the *Sma*I-M5 type plus 2 isolates that were closely and clonal related by *Sma*I typing.

Epiglottitis

- E2 was indistinguishable from E14 when typed using *Sma*I. *Apa*I typing distinguished them and indicated a close or clonal relationship.
- By *Sma*I typing E9 was indistinguishable from E10. *Apa*I typing also showed E9 and E10 to be indistinguishable and included E6, E12, and E14 in this type. As noted above E2 was indistinguishable from E14 by *Sma*I typing but it was shown to be clonally related to it by *Apa*I typing. E9-E14 and E9-E12 were clonally related by *Sma*I typing. But the *Sma*I *F* value of both E6-E9 and E6-E10 was 0.76 suggesting that they may not be closely related. *F* values >0.8 were found between E6-E12 and E6-E14.

Combined Sample

- M15, E5, and E11 were indistinguishable by *Apa*I typing and fell into three different *Sma*I types that were closely or clonally related.
- M20 and E15, indistinguishable by *Apa*I typing, were different by *Sma*I typing but were found to have a clonal relationship.

...text follows Figures 3.7 – 3.15 and Table 3.6

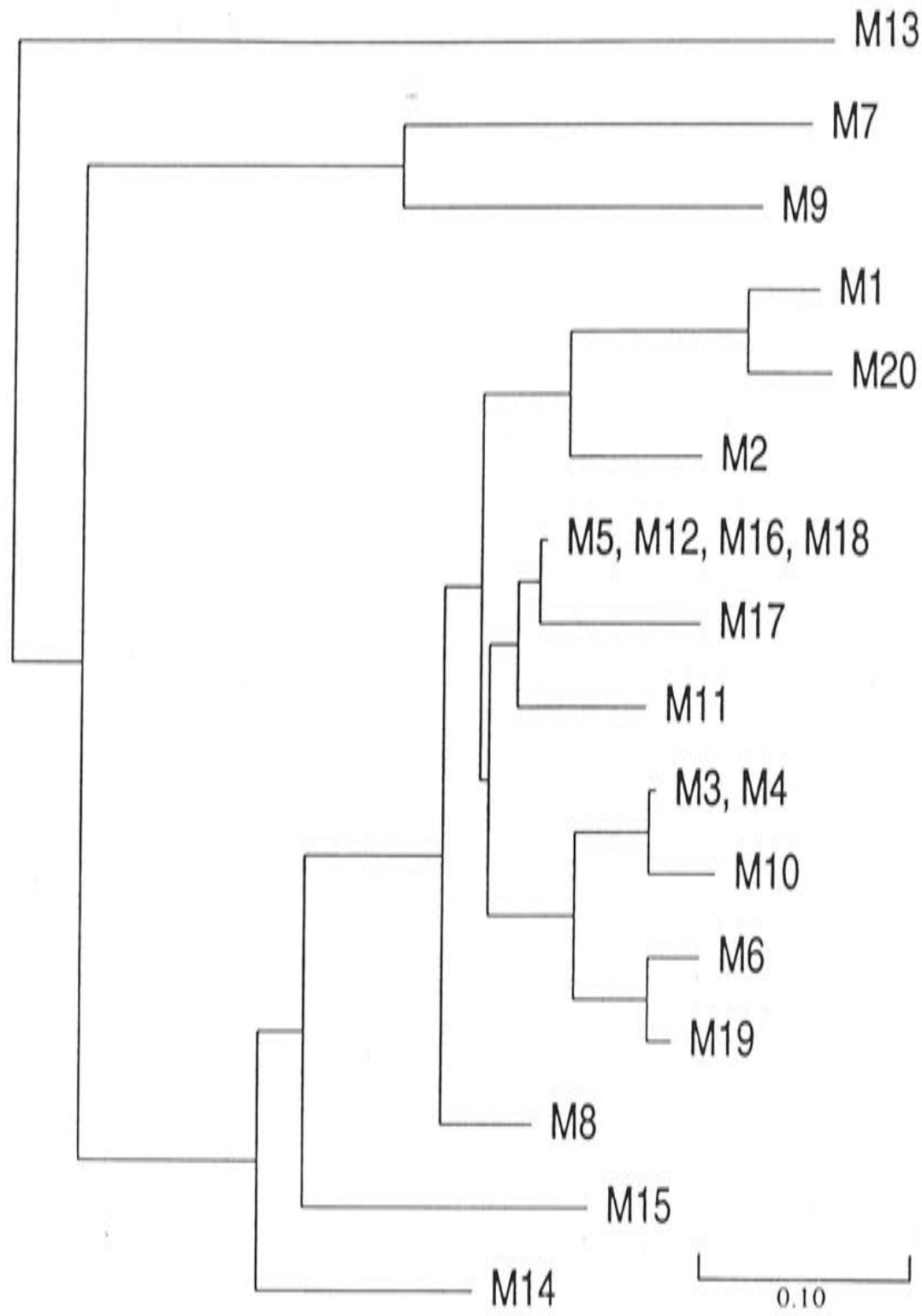


Figure 3.7 **Dendrogram showing the clustering of 16 *Sma*I-RFLP types found among 20 meningitis isolates obtained from patients in the Canberra region.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the length of the horizontal lines between pairs represents the genetic distance between the two types.

Types that fell in clusters ($n=3$) at $F \leq 0.5$	Types that fell in the clusters or branches ($n=12$) at $F \geq 0.9$		Pairs of types ($n=13$) at $F \geq 0.9$	
1) M13	1) M13	7) M6, M19	M1-M20	M5-M8
2) M7, M9	2) M7	8) M3, M10	M2-M5	M5-M11
3) All the other types	3) M9	9) M5, M11	M3-M5	M5-M17
	4) M14	10) M17	M3-M6	M5-M19
	5) M15	11) M2	M3-M10	M6-M19
	6) M8	12) M1, M20	M3-M11	M10-M19
			M3-M19	

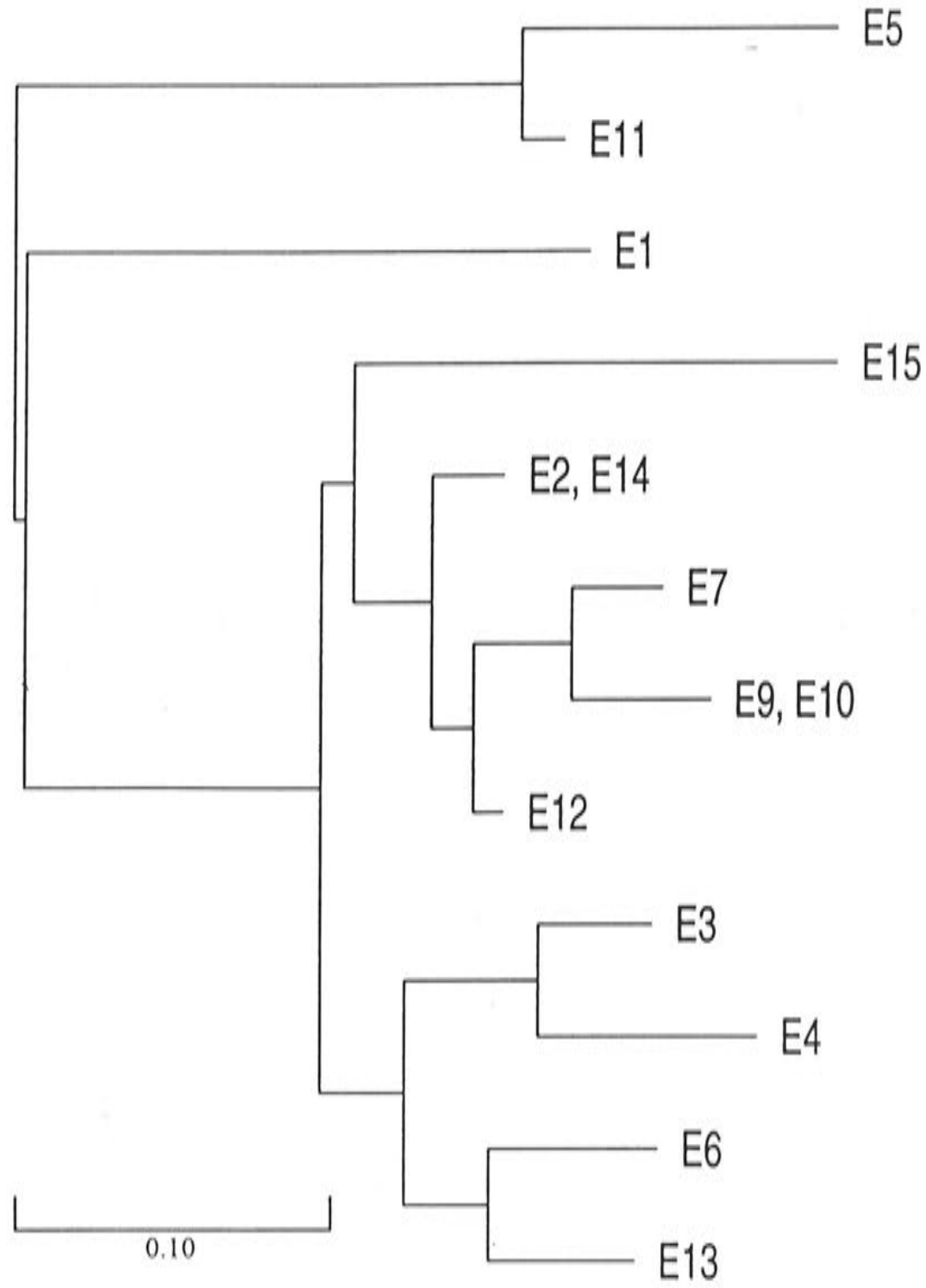


Figure 3.8 **Dendrogram showing the clustering of 12 *Sma*I-RFLP types found among 14 epiglottitis isolates obtained from patients in the Canberra region.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the horizontal lines between pairs is an estimate of the genetic distance between the two types.

Types that fell in clusters ($n=2$) at $F \leq 0.5$	Types that fell in the clusters or branches ($n=6$) at $F \geq 0.9$	Pairs of types ($n=9$) at $F \geq 0.9$	
1) E5, E11	1) E5, E11	E2-E7	E7-E9
2) All the other types	2) E1	E2-E9	E7-E12
	3) E15	E2-E12	E9-E12
	4) E2, E7, E9, E12[a]	E3-E4	
	5) E3, E4	E5-E11	
	6) E6, E13	E6-E13	

[a] This type is indistinguishable from RFLP type *Sma*I-M5.

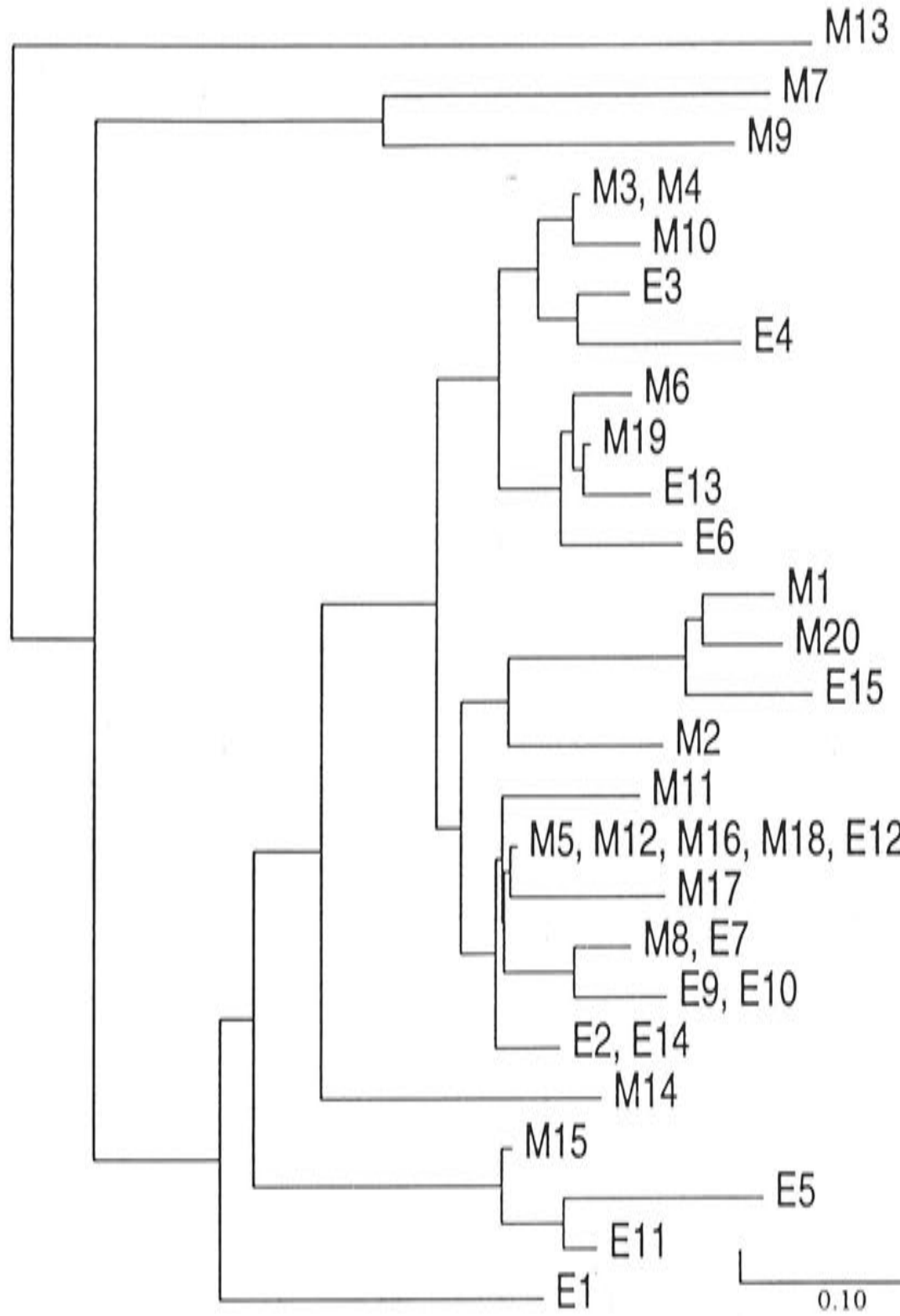


Figure 3.9 **Dendrogram showing the clustering of 26 *Sma*I-RFLP types found among 34 meningitis (*n=20*) and epiglottitis (*n=14*) isolates obtained from patients in the Canberra region.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the horizontal lines between pairs is an estimate of the genetic distance between the two types.

Types that fell in clusters (*n=3*) at $F \leq 0.5$	Types that fell in the clusters or branches (*n=14*) at $F \geq 0.9$		Pairs of types (*n=34*) at $F \geq 0.9$		
1) M13	1) M13	13) M15, E5, E11	M1-M20	M5-M17	M15-E11
2) M7, M9	2) M7	14) E1	M1-E15	M5-M19	M17-E2
3) All the other types	3) M9		M2-M5	M5-E2	M19-E3
	4) M3, M10, E3		M3-M5	M5-E9	M19-E6
	5) E4		M3-M6	M6-M19	M19-E13
	6) M6, M19, E6, E13		M3-M10	M6-E6	M20-E15
	7) M1, M2, M15		M3-M11	M6-E13	E2-E9
	8) M2		M3-M19	M8-E2	E3-E4
	9) M11, M5, M17		M3-E3	M8-E9	E5-E11
	10) M17		M3-E13	M10-M19	E6-E13
	11) M8, E9, E2		M5-M8	M10-E3	
	12) M14		M5-M11	M11-E2	

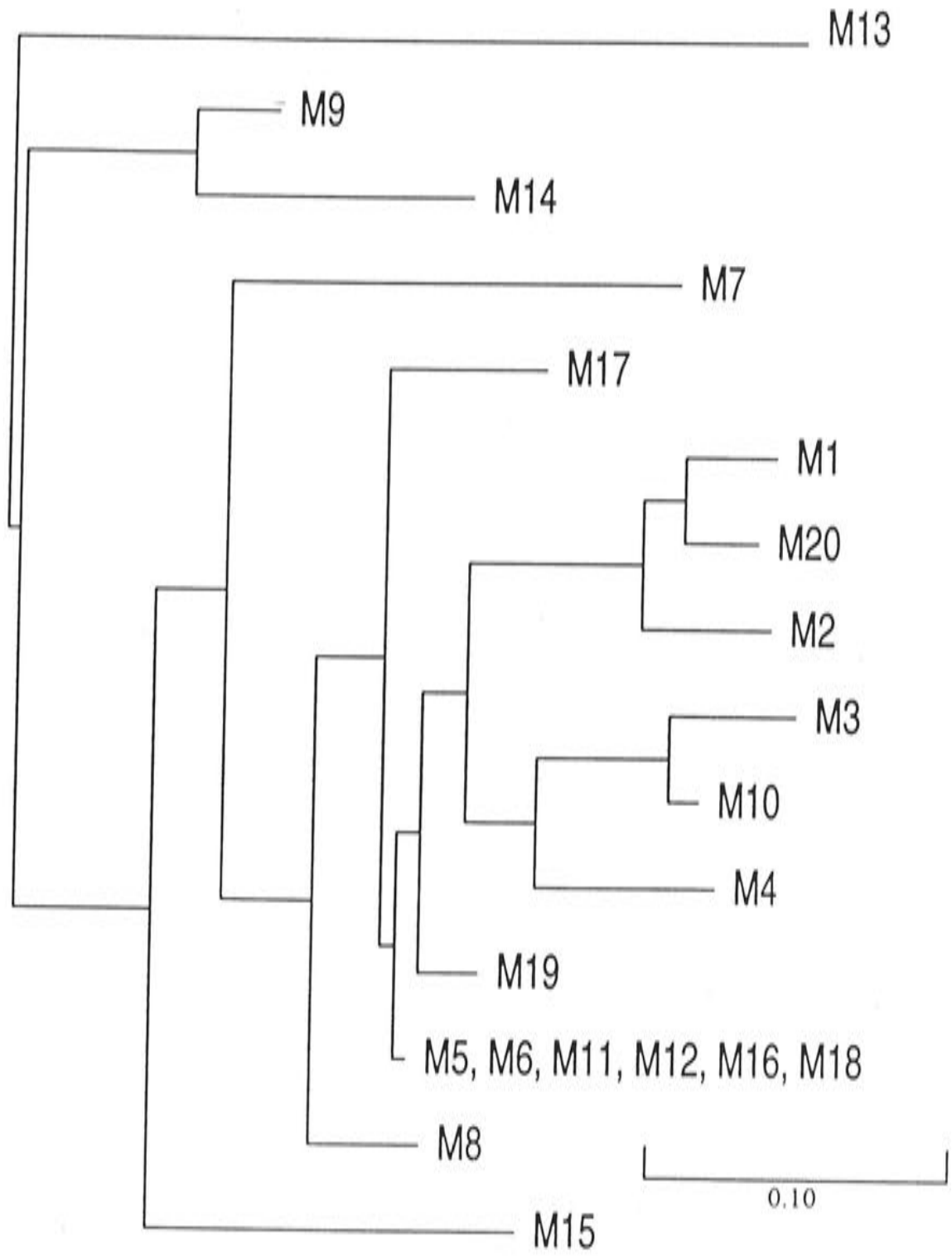


Figure 3.10 **Dendrogram showing the clustering of 15 *Apa*I-RFLP types found among 20 meningitis isolates obtained from patients in the Canberra region.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the horizontal lines between pairs is an estimate of the genetic distance between the two types.

Types that fell in clusters ($n=2$) at $F \leq 0.5$	Types that fell in the clusters or branches ($n=7$) at $F \geq 0.9$	Pairs of types ($n=14$) at $F \geq 0.9$	
1) M13	1) M13	M1-M2	M5-M10
2) All the other types	2) M14	M1-M20	M5-M17
	3) M9	M2-M20	M5-M19
	4) M15	M3-M10	M8-M19
	5) M5, M8, M17, M19	M4-M5	M10-M19
	6) M3, M10, M4	M4-M10	M17-M19
	7) M1, M2, M20	M5-M8	M19-M20

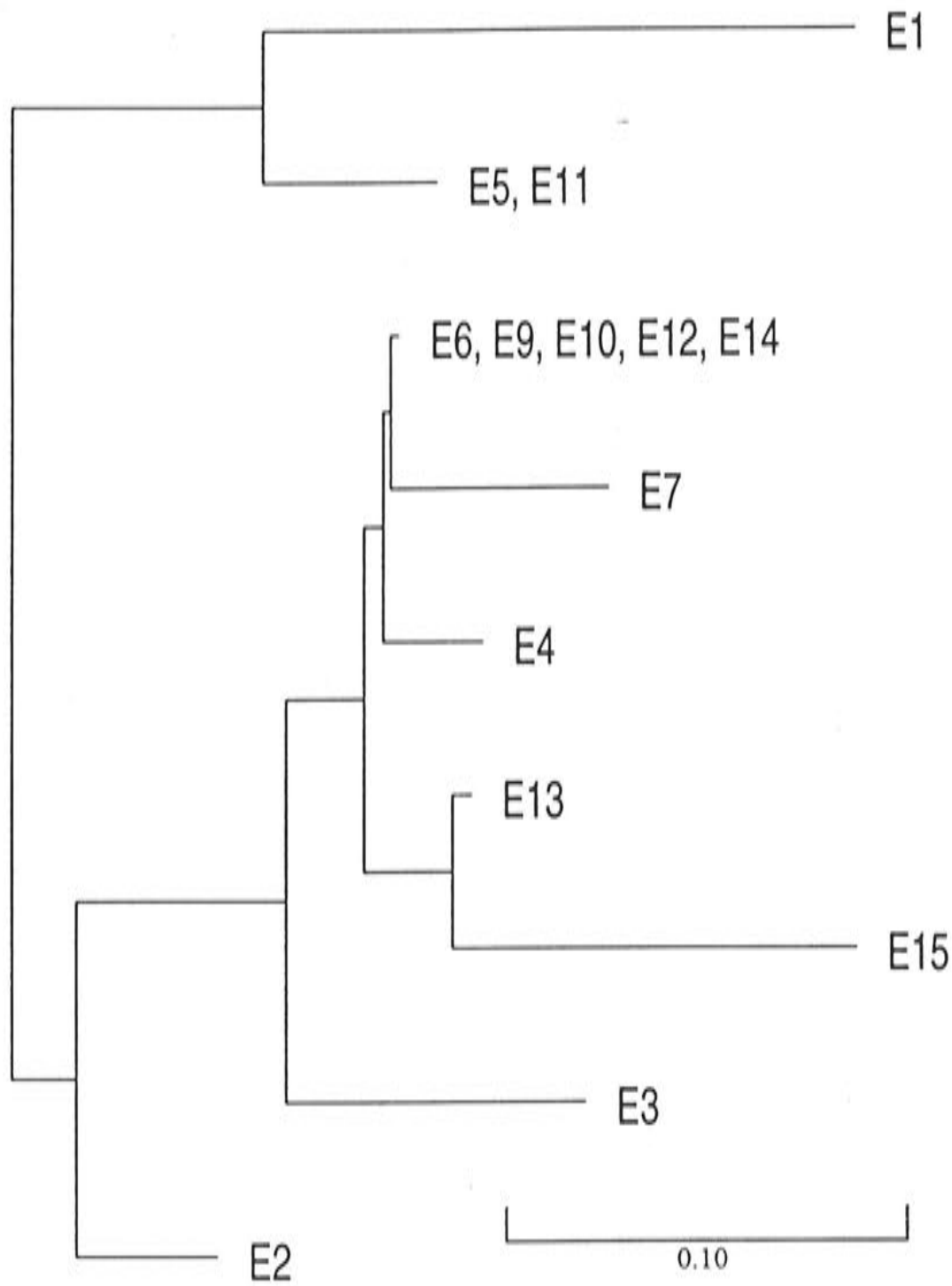


Figure 3.11 **Dendrogram showing the clustering of 9 *Apa*I RFLP types found among 14 epiglottitis isolates obtained from patients in the Canberra region.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the horizontal lines between pairs is an estimate of the genetic distance between the two types.

Types that fell in clusters at $F \leq 0.5$	Types that fell in the clusters or branches ($n=6$) at $F \geq 0.9$	Pairs of types ($n=8$) at $F \geq 0.9$
All the types fell in a single cluster	1) E1 2) E5 3) E4, E6[a], E7, E13 4) E15 5) E3 6) E2	E3-E6 E4-E6 E4-E7 E4-E13 E6-E7 E6-E13 E7-E13 E13-E15

[a] This type is indistinguishable from RFLP type *Apa*I-M5.

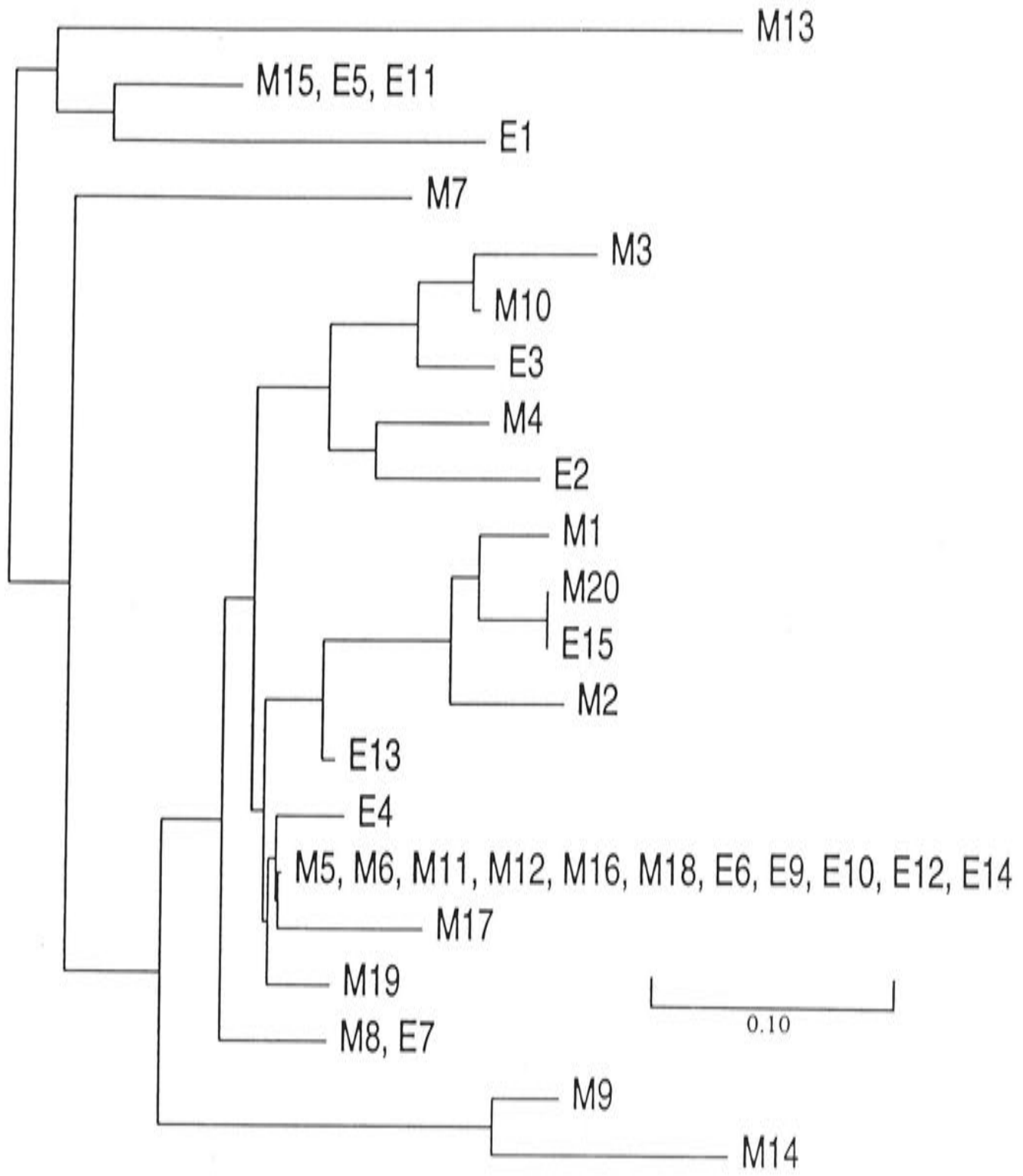


Figure 3.12 **Dendrogram showing the clustering of 20 *Apa*I-RFLP types found among 34 meningitis (*n=20*) and epiglottitis (*n=14*) isolates obtained from patients in the Canberra region.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the horizontal lines between pairs is an estimate of the genetic distance between the two types.

Types that fell in clusters (*n=2*) at $F \leq 0.5$	Types that fell in the clusters or branches (*n=11*) at $F \geq 0.9$	Pairs of types (*n=31*) at $F \geq 0.9$		
1) M13	1) M13	M1-M2	M5-M8	M10-E3
2) All the other types	2) M15	M1-M20	M5-M10	M17-M19
	3) E1	M1-E13	M5-M17	M17-E4
	4) M7	M2-M20	M5-M19	M17-E13
	5) M14	M2-E13	M5-E3	M19-M20
	6) M9	M3-M10	M5-E4	M19-E4
	7) M8	M3-E3	M5-E13	M19-E13
	8) M5, M17, M19, E4, E13	M4-M5	M8-M19	M20-E13
	9) M1, M2, M20[a]	M4-M10	M8-E4	E4-E13
	10) M4, E2	M4-E2	M8-E13	
	11) M3, M10, E3	M4-E3	M10-M19	

[a]This *Apa*I-RFLP type included isolates M20 and E15.

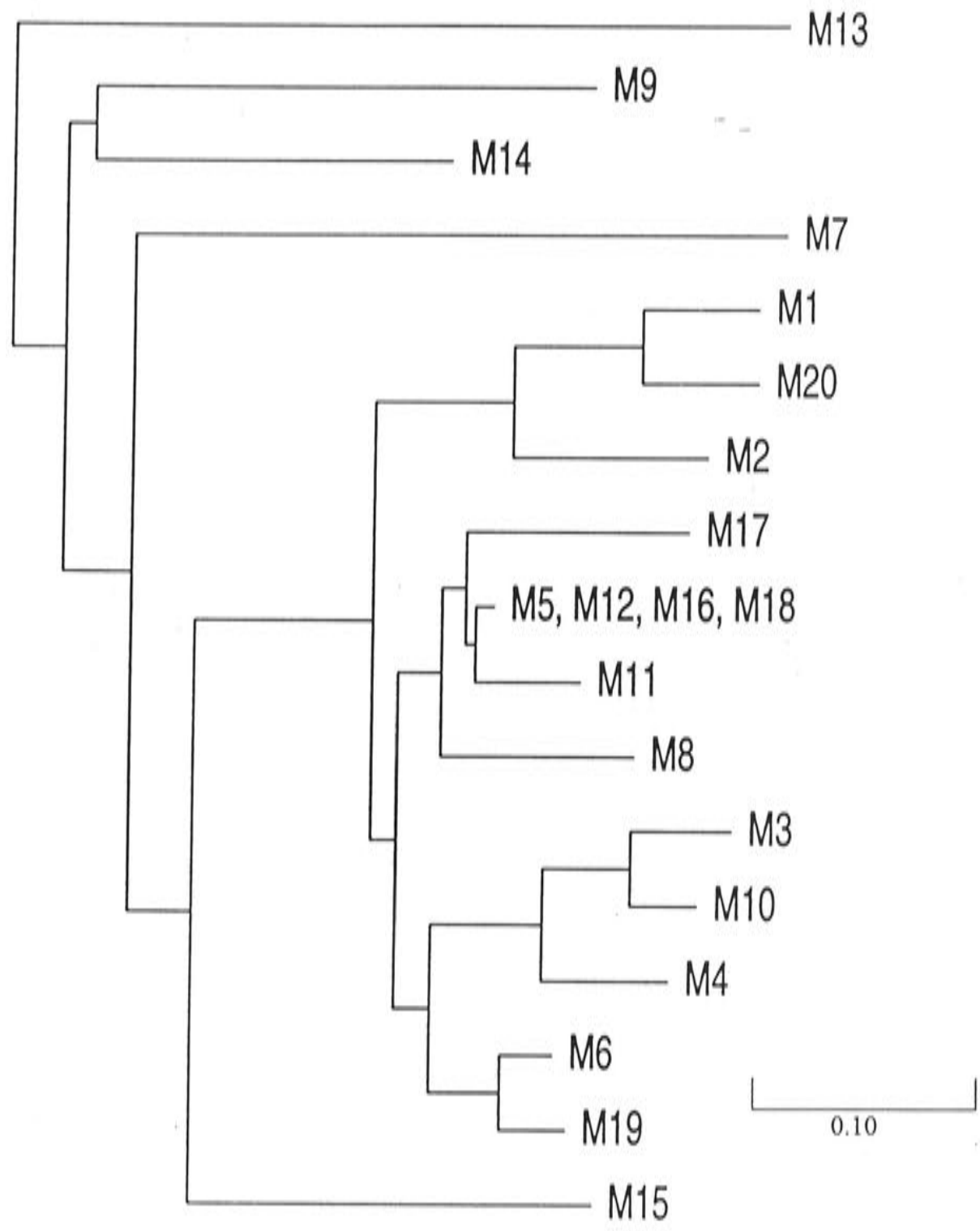


Figure 3.13 **Dendrogram showing the clustering of 17 *Sma*I/*Apa*I-RFLP types found among 20 meningitis isolates obtained from patients in the Canberra region.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the horizontal lines between pairs is an estimate of the genetic distance between the two types.

Types that fell in clusters ($n=2$) at $F \leq 0.5$	Types that fell in the clusters or branches ($n=10$) at $F \geq 0.9$	Pairs of types ($n=22$) at $F \geq 0.9$	
1) M13	1) M13	M1-M2	M5-M8
2) All the other types	2) M9	M1-M20	M5-M11
	3) M14	M3-M4	M5-M17
	4) M7	M3-M10	M5-M19
	5) M15	M3-M19	M6-M8
	6) M8	M4-M5	M6-M11
	7) M5, M11, M17	M4-M6	M6-M19
	8) M6, M19	M4-M10	M8-M11
	9) M3, M4, M10	M4-M11	M10-M19
	10) M1, M2, M20	M4-M19	M11-M17
		M5-M6	M11-M19

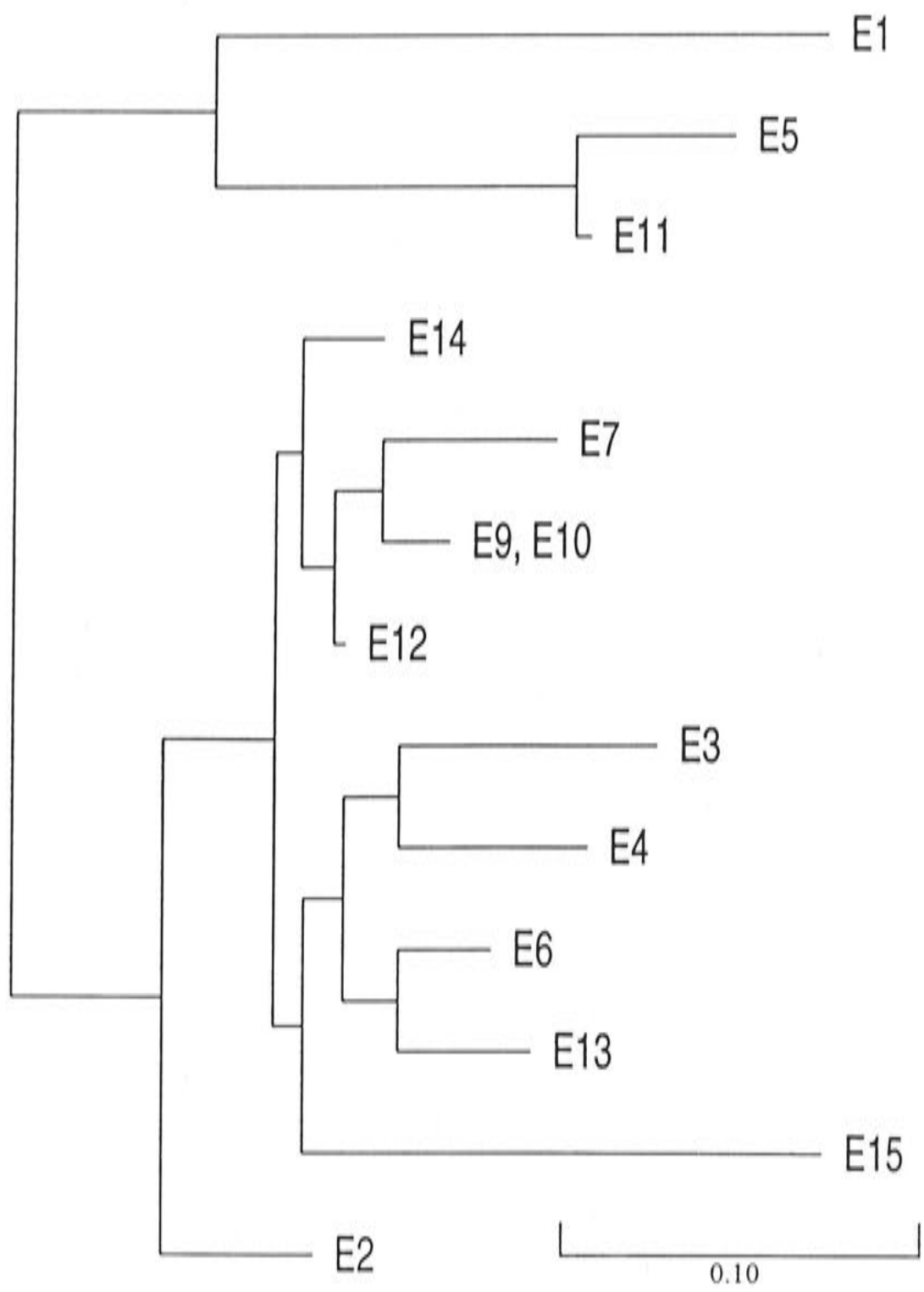


Figure 3.14 **Dendrogram showing the clustering of 13 *Sma*I/*Apa*I-RFLP types found among 14 epiglottitis isolates obtained from patients in the Canberra region.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the length of the horizontal lines between pairs represents the genetic distance between the two types.

Types that fell in clusters at $F \leq 0.5$	Types that fell in the clusters or branches ($n=7$) at $F \geq 0.9$	Pairs of types ($n=20$) at $F \geq 0.9$	
All the types fell in a single cluster	1) E1	E2-E12	E6-E14
	2) E5, E11	E2-E14	E7-E9
	3) E2	E4-E6	E7-E12
	4) E7, E9, E12[b], E14	E4-E9	E7-E14
	5) E4, E6, E13	E4-E12	E9-E12
	6) E3	E4-E13	E9-E13
	7) E15	E5-E11	E9-E14
		E6-E9	E12-E13
		E6-E12	E12-E14
		E6-E13	E13-E14

[a]This type was indistinguishable from RFLP type *Sma*I/*Apa*I-M5.

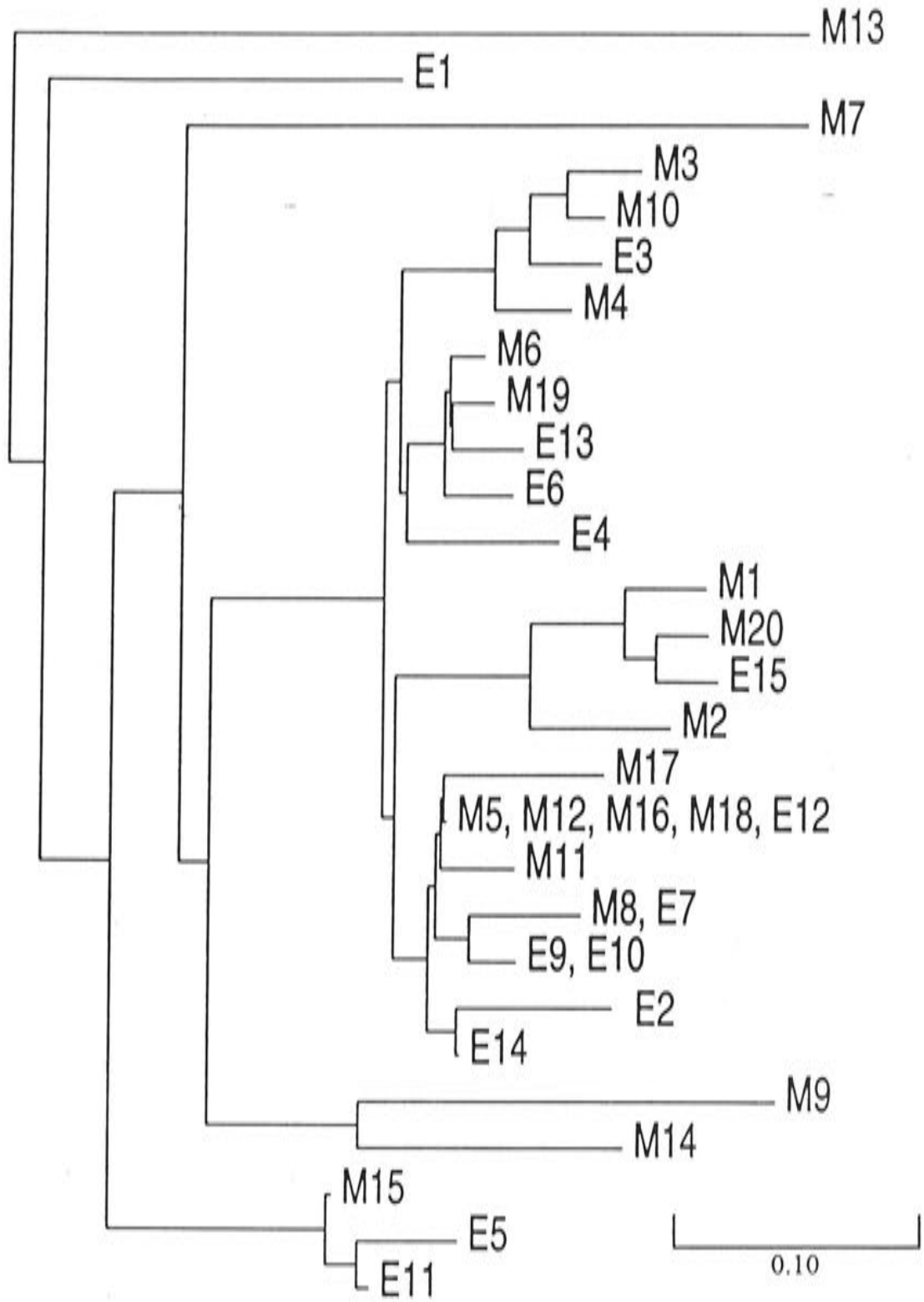


Figure 3.15 **Dendrogram showing the clustering of 28 *Sma*I/*Apa*I-RFLP types found among 34 meningitis (*n=20*) and epiglottitis (*n=14*) isolates obtained from patients in the Canberra region.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the horizontal lines between pairs is an estimate of the genetic distance between the two types.

Types that fell in clusters (*n=2*) at $F \leq 0.5$	Types that fell in the clusters or branches (*n=10*) at $F \geq 0.9$
1) M13	1) M13
2) All the other types	2) E1
	3) M15, E5, E11
	4) M7
	5) M4, M9
	6) M5, M8, M11, M17, E2, E9, E14
	7) M2
	8) M1, M20, E15
	9) M6, M19, E4, E6, E13
	10) M3, M4, M10, E3

Pairs of types (*n=65*) at $F \geq 0.9$

M1-M2	M5-M6	M6-M19	M11-E6	M20-E15
M1-M20	M5-M8	M6-E4	M11-E9	E2-E14
M1-E15	M5-M11	M6-E6	M11-E13	E4-E6
M3-M4	M5-M17	M6-E9	M11-E14	E4-E9
M3-M10	M5-M19	M6-E13	M15-E5	E4-E13
M3-M19	M5-E2	M6-E14	M15-E11	E4-E14
M3-E3	M5-E4	M8-M11	M17-E9	E5-E11
M4-M5	M5-E6	M8-E9	M17-E14	E6-E9
M4-M6	M5-E9	M8-E14	M19-E4	E6-E13
M4-M10	M5-E13	M10-M19	M19-E6	E6-E14
M4-M11	M5-E14	M10-E3	M19-E9	E9-E13
M4-M19	M6-M8	M11-M17	M19-E13	E9-E14
M4-E3	M6-M11	M11-M19	M19-E14	E13-E14

Table 3.6 Summary of the PFGE analysis of the genetic relationships among the meningitis (*n=20*) and epiglottitis (*n=14*) isolates recovered from patients in the Canberra region using the restriction endonucleases, *Sma*I and *Apa*I, and Nei and Li's mathematical model

	Meningitis			Epiglottitis			Meningitis and Epiglottitis		
	*Sma*I	*Apa*I	*Sma*I/*Apa*I	*Sma*I	*Apa*I	*Sma*I/*Apa*I	*Sma*I	*Apa*I	*Sma*I/*Apa*I
Number of RFLP types[a]	16	15	17	12	9	13	26	20	28
Ratio of types to isolates[b]	0.80	0.75	0.85	0.86	0.64	0.93	0.76	0.59	0.82
F value of the most closely related pairs of the different types[c]	0.97	0.97	0.98	0.96	0.97	0.97	0.97	0.97	0.98
F value of the most diverse genetic relationship among pairs of types[d]	0.19	0.43	0.36	0.48	0.53	0.55	0.19	0.43	0.36
Number of shared fragments among the most diverse pair of isolates and the total number of fragments among the pair[e]	3 of 28	8 of 29	12 of 54	7 of 22	9 of 25	18 of 47	3 of 28	8 of 29	12 of 54
Number of clusters among types at $F \leq 0.5$[f]	3	2	2	2	1	1	3	2	2
Number of clusters among types at $F \geq 0.9$[g]	12	7	10	6	6	7	14	11	10
Number of pairs among types at $F \geq 0.9$[h]	13	14	22	9	8	20	34	31	65

[a] *Sma*I discrimination yielded more RFLP types than *Apa*I.
[b] The ratio of *Sma*I types to isolates is highest and the difference between using *Sma*I alone versus *Sma*I and *Apa*I is not great.
[c] Closely related isolates were consistently identified by each of the enzymes.
[d] An inconsistency in genetic diversity among the most diverse pairs of types among the meningitis isolates was found between *Sma*I and *Apa*I.
[e] This data correlates with the *F* value of the most diverse genetic relationship among pairs of types listed in the previous row.
[f] *Sma*I discrimination produced more clusters at $F \leq 0.5$ than *Apa*I.
[g] *Sma*I discrimination produced more clusters at $F \geq 0.9$ than *Apa*I.
[h] *Sma*I discrimination produced more pairs among types at $F \geq 0.9$ than *Apa*I.

3.4 Discussion

It has been shown that the Hib population is clonal, which implies that the number of different types will be limited (Musser *et al.*, 1988). One practical implication of this is that, in general, non-specific procedures such as RFLP analysis can be used to identify bacteria in such populations (Porras *et al.*, 1986). The aim of this study was to evaluate the use of PFGE for determining the genetic diversity of *Haemophilus influenzae* type b and to characterise a sample of clinical isolates of Hib. Because of its epidemiological significance a sample of 34 isolates from meningitis and epiglottitis patients from Canberra, a discrete geographical area, was selected for this study.

The rare-cutting endonucleases, *Sma*I and *Apa*I, were chosen to digest the Hib DNAs, and 20 *Apa*I-RFLPs, 26 *Sma*I-RFLPs and 28 *Sma*I/*Apa*I-RFLPs were found. Identical, closely related, and distantly related isolates of Hib were distinguished and the correlation between the enzymes was very good. Both enzymes produced distinct, comparable patterns though *Sma*I patterns were easier to interpret than *Apa*I patterns. The average number of fragments per type ranged from 14 to 18, for *Sma*I and *Apa*I digests, respectively. This is a sufficient number for good discrimination with patterns produced that are not excessively complex. Repeat runs were necessary to confirm some observations particularly when comparing isolates on different gels. When necessary, the isolates in question were loaded into adjacent lanes and run together on a new gel. While the need to prepare repeat runs may not ever be entirely avoided, a high level of consistency in producing quality gels is necessary to facilitate comparison of patterns and to limit repeat runs.

Correlation between the two enzymes was very good ($r = 0.78$ with 561 data) but, not surprisingly, discrepancies were found. However, among types with multiple isolates typing differences were mostly due to pairs that were indistinguishable by *Apa*I typing but distinguishable as clonally or closely related when typed with *Sma*I. The most diverse *Sma*I *F* value found between isolates that were indistinguishable using *Apa*I was 0.76 when 7 of the 19 different *Sma*I fragments found between E6-E9 were not shared. Even this difference, however, is not outside the bounds of potentially clonally related isolates (Thal *et al.*, 1997). Thus, among types with multiple isolates, differences in typing between the two enzymes did not change the overall genetic diversity inferred from the results.

Among the RAPDistance Programs package is a program that identifies samples producing patterns that correlate worst with a distance matrix. When it was applied to the distance matrix of all the types, removal of up to 16 isolates was required to increase the correlation coefficient by just 0.04. After removing those 16 isolates no improvement in the correlation coefficient was attained. This suggests that no one isolate was responsible for decreasing the correlation of the enzymes and that typing differences between the two enzymes did not obscure the relationship of the isolates determined by PFGE using either enzyme.

Calculation of Simpson's index of diversity (D) showed that *Sma*I digestion was more discriminating than *Apa*I digestion. D = 0.974 for *Sma*I, indicates a very high power of discrimination. The D value of *Apa*I, 0.890, is considered to be less desirable but is just within the range where results can be interpreted with confidence. The use of both enzymes together did not improve the power of discrimination significantly. Ideally, typing schemes should not be validated with small samples and a larger sample ($n \geq 100$) should be used to confirm the D values. However, one would predict from the number of types found and the presence of multiple isolates per type that this typing system has a high level of discrimination.

It was shown that *Sma*I digestion results were similar to those found using *Apa*I digestion and that both are suitable for typing Hib isolates. The practical implication of this is that either of the enzymes alone could be used to examine the overall genetic diversity of Hib and its population structure with similar results. Since *Sma*I patterns had fewer fragments, were less complex and were easier to compare, I prefer its use from a technical point of view. In addition, *Sma*I was more sensitive in detecting differences among closely related isolates. While the combination of patterns obtained with both enzymes will theoretically increase the power of a typing system, only two additional types were found by *Sma*I/*Apa*I typing versus typing with *Sma*I alone. In each case indistinguishable isolates that were found to be different by the second enzyme were clonally related. Interestingly, a close genetic relationship determined by *Sma*I typing was confirmed by *Apa*I typing among 78 Hib strains recovered from the cerebrospinal fluid of children with meningitis in Italy (Tarasi *et al.*, 1998). The data were not subject to cluster analysis. Estimates of genetic distance were based on the number of shared fragments between pairs without subjecting the data to numerical

analysis. The authors concluded that a combination of the enzymes would enhance the discriminatory power of PFGE but it was not clear to what extent this occurred.

The dendrograms generated from the similarity matrices of each enzyme reflect the consistency of the overall relationships between types. The closely related isolates are consistently closely related and, while the more distantly related isolates tend to vary slightly in their location on the dendrograms, the overall close and distant relationships are not obscured. The entire sample typed by both enzymes was analysed for the effect of the fragment data on the numerical calculation used to determine the genetic distance values. The same clustering and distribution of types found using Nei and Li's *F* value was consistently observed when dendrograms were produced from the similarity matrices calculated using 15 different algorithms (RAPDistance Programs).

The distribution of RFLP types indicates that the population of Hib in Canberra is clonal. None of the types was associated with beta-lactamase production, biotype, age, sex or disease state. The close genetic relationship among most of the isolates digested by *Sma*I was confirmed following *Apa*I digestion. One type, *Sma*I/*Apa*I-M5, and isolates clonally related to it accounted for almost 50% of the sample representing a clone endemic in the Canberra region. These data are in accordance with the results obtained with multilocus enzyme electrophoresis suggesting that *H. influenzae* is basically clonal and that most diseases are caused by a very small number of clonal genotypes (Musser *et al.*, 1990).

The isolates in the predominant clone were almost evenly distributed among the epiglottitis and meningitis isolates. And, isolates from each disease group were found in 6 of the 10 clusters at a similarity ≥0.90 and in the two clusters at a genetic distance of ≤0.50. It may well be that the failure to demonstrate clonal disease specificity was caused by the choice of enzyme. Other restriction enzymes that sample different areas of the genome may detect molecular differences that demonstrate clonal disease specificity but we did not detect such a relationship among the isolates in this study. It can also be argued that owing to the small sample size, more isolates must be studied to confirm the lack of association of PFGE types with epiglottitis.

The epiglottitis isolates were less diverse than the meningitis isolates. The smallest *F* value (0.48) among them was 0.29 greater than the smallest value found among the

meningitis isolates (0.19). However, if isolate E8 had been kept in the sample the diversity of the epiglottitis isolates would have been similar to that of the meningitis isolates because E8 was clonally related to M13. M13 was the most diverse isolate in the sample and genetically distinct from all the other isolates.

An interesting note is that the pattern of RFLP type *Sma*I-M5 was indistinguishable from the *Sma*I pattern of a known OMP subtype 1 strain that we hold in a set of OMP reference isolates. OMP subtype 1 is the predominant European outer membrane protein subtype described by (van Alphen *et al.*, 1983). It has been reported to be identical with that designated 3L among the North American isolates studied in the development of the system reported by Barenkamp et al. (Barenkamp *et al.*, 1981). This subtype accounted for 83% of 187 isolates recovered from children with invasive disease in Victoria (Clements *et al.*, 1992). It was also found to account for 76% of 59 Hib isolates from Adelaide and one from Newcastle (Hansman and Lawrence, 1993).

Furthermore, the profile of *Sma*I-M5 appears to be indistinguishable from the predominant pattern described among Italian isolates that matched a strain from the United Kingdom and was closely related to reference strain ATCC 10211 (Tarasi *et al.*, 1998). Computer generated graphic patterns of *Sma*I-RFLP types of Hib recently reported among Japanese isolates were difficult to compare with the gel pictures from this study but it did not appear that any of those patterns matched *Sma*I-M5 (Mitsuda *et al.*, 1999).

This study showed that PFGE can be used to explore the genetic diversity of Hib's population structure in Australia and has extended our knowledge on the range of its clonality. It showed that similar strains of Hib are associated with meningitis and epiglottitis and that a single type and its clonally related isolates predominate in the Canberra region. The correlation of the predominant RFLP type with OMP subtype 1 suggests that this type may be predominant in other areas of Australia and that it may be of European origin. Further investigation using PFGE to study the diversity of Hib types, the range of the predominant clonal group, and the association of types with disease states is warranted.

CHAPTER 4

Characterisation of 104 Non-Aboriginal Meningitis and Epiglottitis Isolates of *Haemophilus influenzae* type b Using Pulsed Field Gel Electrophoresis

4.1 Introduction

The results described in Chapter 3 showed that PFGE would be useful for characterising genetic types among Hib isolates. To further investigate the variation of Hib types and the lack of association of types with meningitis and epiglottitis found in the preliminary study, meningitis and epiglottitis isolates from three other metropolitan areas were obtained for characterisation by PFGE. The isolates were recovered before the introduction of Hib vaccines from non-Aboriginal patients living in the urban areas of Melbourne, Sydney, and Perth.

It has been shown that the incidence of epiglottitis in Victoria (Gilbert *et al.*, 1990), the Australian Capital Territory (McGregor *et al.*, 1992), and Sydney (McIntyre *et al.*, 1991) was almost twice that found in Western Australia (Hanna *et al.*, 1992). An age difference was also found with 8 and 25% of the Western Australian and Victorian cases, respectively, occurring after 60 months of age (Hanna, 1990). The reason for the differing epidemiology is not known, but, it might be explained if there were differences in the strains of Hib isolated in the different regions.

Unlike that of epiglottitis, the incidence of meningitis in children under the age of 5 years was similar in these four cities. The annual incidence averaged around 25 cases/100,000 (the range was 20-31 cases). The lowest incidence was found in Sydney and thought to be due to under-reporting (McIntyre *et al.*, 1991). In addition, reports from South Australia (El Saadi and Cameron, 1993) and Queensland (Hanna and Wild, 1991) showed similar rates of Hib meningitis though the incidence was slightly higher in Queensland.

Isolates from Aboriginals were not found among the Canberra, Melbourne and Sydney Hib collections that we accessed. Thus, this sample was limited to non-Aboriginal isolates. In Canberra where Aborigines comprise only 0.5% (ABS, 1987) of the population, no patients of Aboriginal descent were identified in a seven year study period (McGregor *et al.*, 1992). Similarly, the population of Victoria is predominantly white and only 0.3% Aboriginal (ABS, 1987). It is not known to what extent isolates from Aboriginal patients are identified in the Sydney metropolitan area but Aboriginality was not mentioned by McIntyre *et al.* (1991) in their study of the epidemiology of Hib disease in this metropolitan region.

This study was undertaken to further investigate the hypotheses that genotypes associated with epiglottitis are not associated with meningitis and vice versa. The lack of association of types with disease state among the 34 isolates in the preliminary study does not substantiate these hypotheses; but, a examination of a larger sample would confirm this. Furthermore, it is not known what types are associated with disease in other areas and whether they are the same as those found in the Canberra region. It would be interesting to determine whether the predominant clone found in the Canberra sample is found elsewhere. The purpose of this study was 1) to ascertain genetic types of Hib associated with different areas and compare them for similarity, and 2) to determine whether any types are associated with meningitis or epiglottitis among isolates sampled from different areas.

4.2 Methods

Preparation of DNA, restriction endonuclease digestion, PFGE conditions, determination of the numbers and mobilities of fragments, estimation of genetic diversity and construction of dendrograms were performed as previously described. Hybridisation, plasmid analysis and routine biochemical tests were also performed as previously described.

4.2.1 Bacterial isolates (Table 4.1, Appendices A.1-A.4)

The 104 isolates included 67 recovered from patients with meningitis and 37 recovered from patients with epiglottitis. All were collected between 1985 and 1990. Thirty-four isolates from Canberra (previously characterised by PFGE and described in Chapter 3) and 22 isolates from Sydney were obtained from the Canberra Hospital, Australian Capital Territory; 20 isolates were obtained from the Royal Children's Hospital, Melbourne, Victoria; and 28 isolates from were obtained from the Princess Margaret Hospital, Perth, Western Australia. A breakdown of the geographic location and disease association of the isolates is shown in Table 4.1. Appendices 1-4 describe the isolates in more detail.

All of the isolates obtained for this study had been identified by conventional methods as described previously. They were stored at –70°C or, in the case of the isolates from Western Australia, were lyophilised. Because of the strong correlation between types using *Sma*I and *Apa*I found in the preliminary study only one enzyme, *Sma*I was used

for the analysis of RFLPs in this study. It was chosen because the number and size of fragments produced using *Sma*I were fewer and more distinct than those of *Apa*I. Digestion with *Apa*I was performed and the results were available for backup if needed.

Table 4. 1 Geographic location and disease association of 104 Hib isolates recovered from non-Aboriginals diagnosed with meningitis or epiglottitis

Geographic location	Number of isolates		
	Meningitis	Epiglottitis	Total
Canberra, Australian Capital Territory[a]	20	14	34
Melbourne, Victoria[b]	10	10	20
Sydney, New South Wales[a]	21	1	22
Perth, Western Australia[c]	16	12	28
Total	67	37	104

[a]From a collection held at Canberra Hospital, Woden, Australian Capital Territory
[b]From a collection held at Royal Children's Hospital, Melbourne, Victoria
[c]From a collection held at Princess Margaret Hospital, Perth, Western Australia

The isolates were designated as follows:

Source	Meningitis	Epiglottitis
Canberra (*n=34*)	M1-20 (*n=20*)	E1-7, E9-15 (*n=14*)[a]
Victorian (*n=20*)	V1-10 (*n=10*)	V11-20 (*n=10*)
Sydney (*n=22*)[b]	S2, 5-9, 11-12, 14, 16-18, 20, 22, 24-25, 28-30, 35-36 (*n=21*)	S37 (*n=1*)
Perth (*n=28*)[b]	WA14-17, 22, 24-25, 27, 29-31, 34, 41, 43-44, 47 (*n=16*)	WA18-21, 26, 28, 35-36, 40, 42, 45-46 (*n=12*)

[a]The isolate designated E8 was excluded from the Canberra subset when it was found to be from Campbelltown, NSW, and subsequently was designated S37
[b]Breaks in the numbering represent Hib isolates in our collection that were recovered from patients diagnosed with diseases other than meningitis or epiglottitis. These 'other' isolates were analysed and included in the study described in Chapter 6.

4.3 Results

4.3.1 Hybridisation

All the isolates hybridised with pU082. Hybridisation was detected on one or two DNA fragments per isolate on the *Sma*I-PFGE gels. For most types the same size fragment, corresponding to fragment 8.1 (~416 kb) of HS008, hybridised with pU082. Among sets of multiple types the same fragment or fragments hybridised and no further splitting of RFLP types was found from the capsular typing of the PFGE RFLPs.

4.3.2 Visual analysis of *Sma*I-RFLPs

Tables 4.2-4.4 describe the fragment distribution among types found in each subset. Fragment sizes ranged from ~9-500 kb which extended the size range found in the preliminary study. The increase in fragment size was due to just two isolates, S29 and S37. Most isolates were tested at least twice.

Canberra isolates (Figures 3.1-3.3, Table 3.1) The analysis of these isolates is described in Chapter 3. In summary, 26 *Sma*I-RFLPs were found among 34 isolates. A predominant type and its clonally related types represented 47% of the sample. Only two types were found among both the meningitis and epiglottitis isolates. The most genetically distant pair of isolates shared just 3 of their 28 fragments (F=0.19).

Victorian isolates (Figures 4.1-4.2, Table 4.2) Six types were found among the 10 meningitis isolates. One accounted for 40% of the sample and was indistinguishable from the predominant type found among the Canberra isolates. One of the four types found among the 10 epiglottitis isolates was also indistinguishable from the predominant Canberra type. It accounted for 7 (70%) of the epiglottitis isolates. The distribution of isolates among types was:

Meningitis isolates (n=10)		Epiglottitis isolates (n=10)	
*Sma*I-RFLP type	Number of isolates/type	*Sma*I-RFLP type	Number of isolates/type
V2 (same as V11)*	4	V11 (same as V2)*	7
V1	2	V14, V18, V19	1 each
V5, V7, V8, V10	1 each		

*indistinguishable from RFLP type *Sma*I-M5 that predominated in the Canberra sample

When the meningitis and epiglottitis isolates were combined 9 types were found. The ratio of types to isolates was lower and the types were less diverse than those found in the Canberra sample. The type that was indistinguishable from the predominant Canberra type accounted for 55% of the sample. It was the only type found among both the meningitis and the epiglottitis isolates. The distribution of isolates among the combined meningitis and epiglottis types was:

Meningitis and epiglottitis isolates combined (*n=20*)

*Sma*I-RFLP type	Number of isolates per type	Distribution of isolates/type/disease	
		Meningitis	Epiglottitis
V2*	11	4	7
V1	2	2	0
all others (*n=7*)	1 each	V5, V7, V8, V10	V14,V18,V19

*indistinguishable from RFLP type *Sma*I-M5 that predominated in the Canberra sample

Sydney isolates (Figures 4.2-4.6, Table 4.3) Eleven types were found among the 21 meningitis isolates from Sydney. One of them was indistinguishable from RFLP type M5, the predominant type found in the Canberra and Victorian samples. It accounted for 41% of the sample. The only epiglottitis isolate in the sample was a unique type and the most diverse one. The distribution of isolates among types was:

Meningitis and epiglottitis isolates combined (*n=22*)

*Sma*I-RFLP type	Number of isolates/type
S6*	9
S16	2
S2, 5, 8, 11, 12, 14, 20, 29, 37	1 each

*indistinguishable from RFLP type *Sma*I-M5 that predominated in the Canberra sample

Western Australian isolates (Figures 4.7-4.9, Table 4.4) Five types were found among the meningitis isolates from Perth. One was indistinguishable from RFLP type M5, the predominant Canberra/Melbourne/Sydney type. It accounted for 7 (44%) of the 16 meningitis isolates and 5 (42%) of the 12 epiglottitis isolates among which 7 types were found. The distribution of isolates among types was:

Meningitis isolates		Epiglottitis isolates	
*Sma*I-RFLP type	Number of isolates/type	*Sma*I-RFLP type	Number of isolates/type
WA14 (same asWA18)*	7	WA18 (same as WA14)*	5
WA15	4	WA19	2
WA16	2	WA20,21,36,42,46	1 each
WA24	2		
WA17	1		

*indistinguishable from RFLP type *Sma*I-M5 that predominated in the Canberra sample

In the combined sample 9 types were found. The predominant type, that was indistinguishable from M5 included 12 (43%) of the 28 isolates. It was clonally related

...text continues after Tables 4.2-4.4 and Figures 4.1-4.9

Table 4. 2 Number of *Sma*I-RFLP types and the number of fragments found per type among 20 meningitis and epiglottitis Hib isolates from Victoria

	Total number of types		Number of fragments per type 13	14	15
Meningitis isolates *n=10*	6 (0.60:1)[a]	Types	3	1	2
		Isolates	*n=6*	*n=1*	*n=3*
Epiglottitis isolates *n=10*	4 (0.40:1)[a]	Types	2	0	2
		Isolates	*n=8*	*n=0*	*n=2*
All isolates *n=20*	9[b] (0.45:1)[a]	Types	4[c]	1	4
		Isolates	*n=14*	*n=1*	*n=5*

[a]Ratio of the number of types found to the total number of isolates typed.
[b]One *Sma*I-RFLP type was found among both meningitis and epiglottitis isolates.
[c]One of the 13 fragment types was found among both meningitis and epiglottitis isolates.

Table 4. 3 Number of *Sma*I-RFLP types and the number of fragments found per type among 22 meningitis and epiglottitis Hib isolates from Sydney, New South Wales

	Total number of types		Number of fragments per type 13	14	15	16
Meningitis isolates *n=21*	11 (0.52:1)[a]	Types	4	6	1	
		Isolates	*n=13*	*n=6*	*n=2*	
Epiglottitis isolates *n=1*	1	Types				1
		Isolates				*n=1*
All isolates *n=22*	12 (0.45:1)[a]	Types	4	6	1	1
		Isolates	*n=13*	*n=6*	*n=2*	*n=1*

[a]Ratio of the number of types found to the total number of isolates typed.

Table 4. 4 Number of *Sma*I-RFLP types and the number of fragments found per type among 28 meningitis and epiglottitis Hib isolates from Western Australia

	Total number of types		Number of fragments per type 13	14
Meningitis isolates *n=16*	5 (0.31:1)[a]	Types	3	2
		Isolates	*n=13*	*n=3*
Epiglottitis isolates *n=12*	7 (0.58:1)[a]	Types	4	3
		Isolates	*n=9*	*n=3*
All isolates *n=28*	9[b] (0.32:1)[a]	Types	5[c]	4[d]
		Isolates	*n=23*	*n=5*

[a]Ratio of the number of types found to the total number of isolates typed.
[b]Three *Sma*I-RFLP types were found among both meningitis and epiglottitis isolates.
[c]Two of the 13 fragment types were found among both meningitis and epiglottitis isolates.
[d]One of the 14 fragment types was found among both meningitis and epiglottitis isolates.

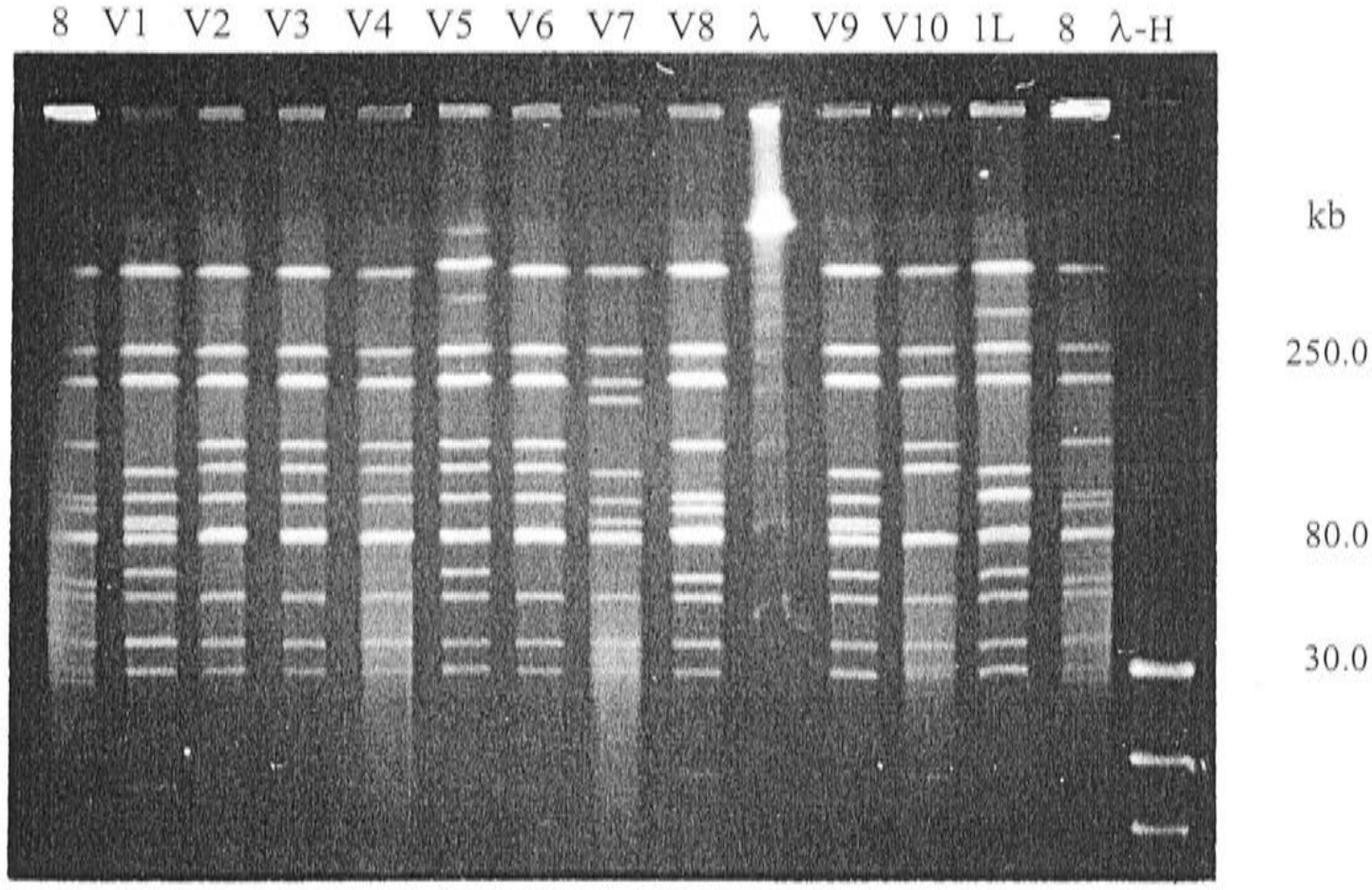


Figure 4.1 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 10 Hib isolates recovered from the cerebrospinal fluid or blood of meningitis patients from Victoria. Lanes labelled 8, λ and λ-H contain fragment standards; lane 1L is a Hib isolate with OMP subtype 1L from a set of OMP standards.

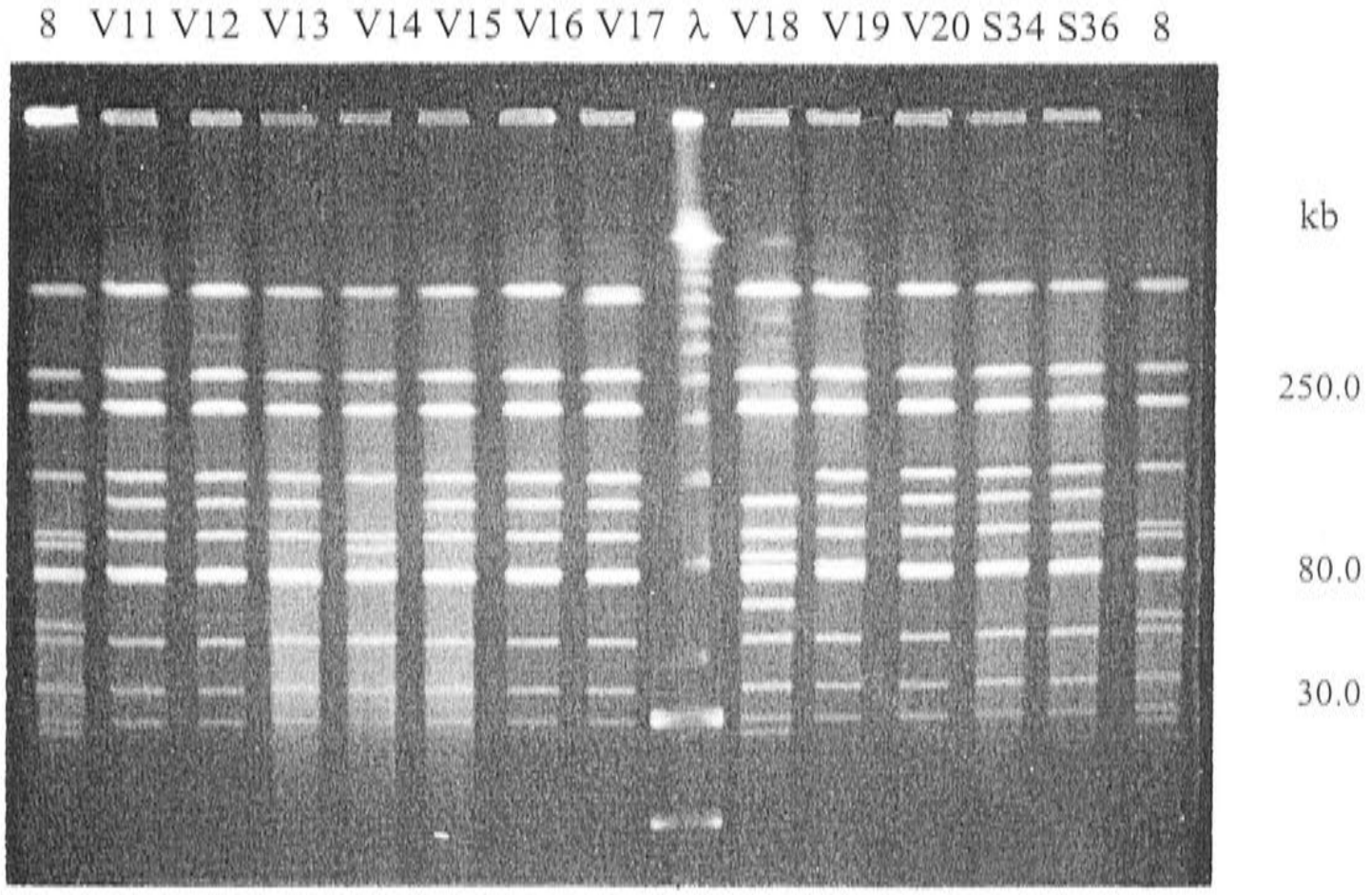


Figure 4.2 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 10 Hib isolates recovered from the blood of epiglottitis patients from Victoria and 2 isolates from a meningitis patient (S36) and a patient with invasive disease from the Sydney region. Lanes labelled 8 and λ contain fragment standards.

Figure 4.3 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 10 Hib isolates recovered from the cerebrospinal fluid or blood of meningitis (m) patients or patients with other (o) infections from the Sydney region. Lanes labelled 8, λ and λ-H contain fragment standards.

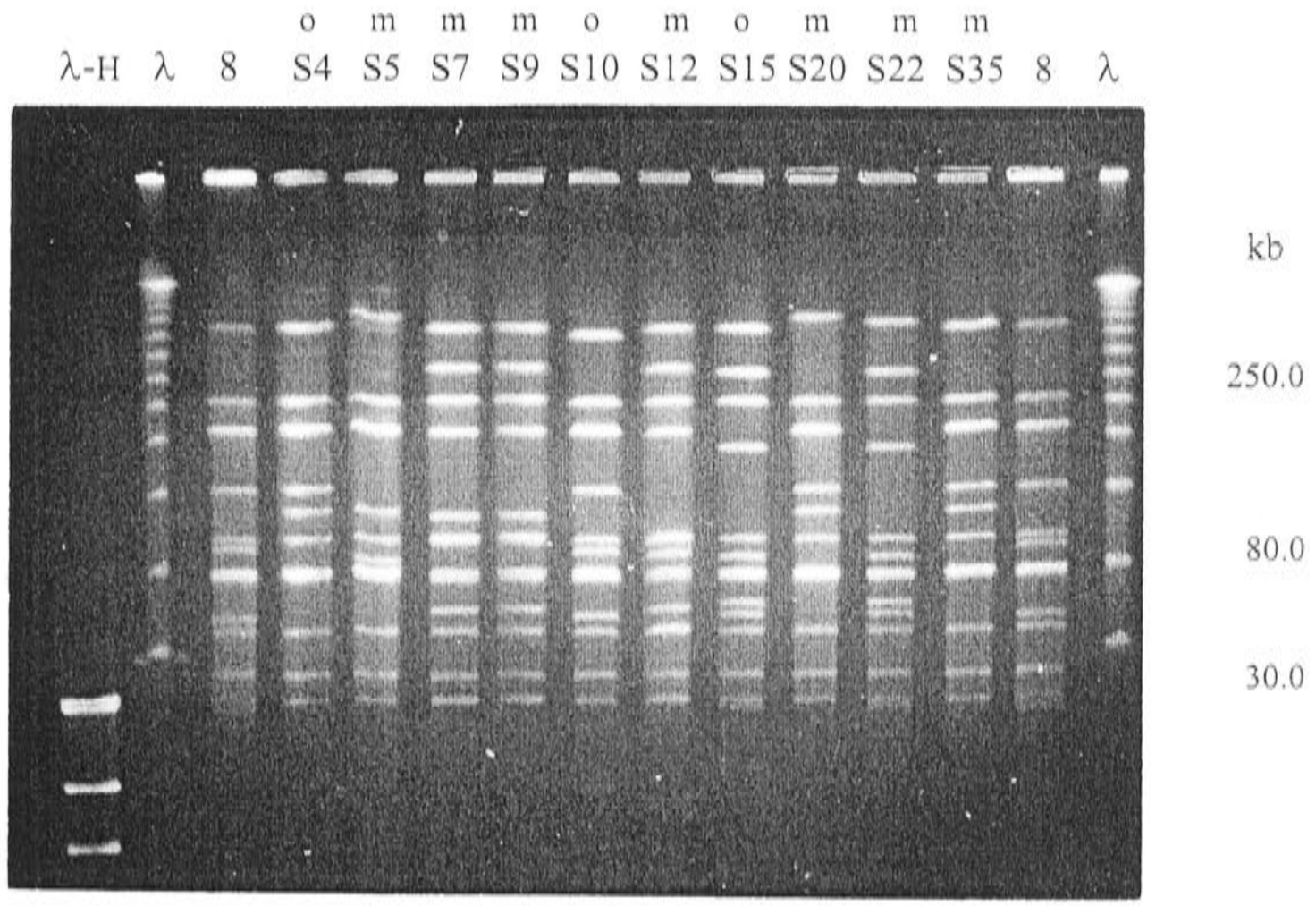


Figure 4.4 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 10 Hib isolates recovered from the cerebrospinal fluid or blood of meningitis (m) patients or patients with other (o) infections from the Sydney region. Lanes labelled 8, λ and λ-H contain fragment standards.

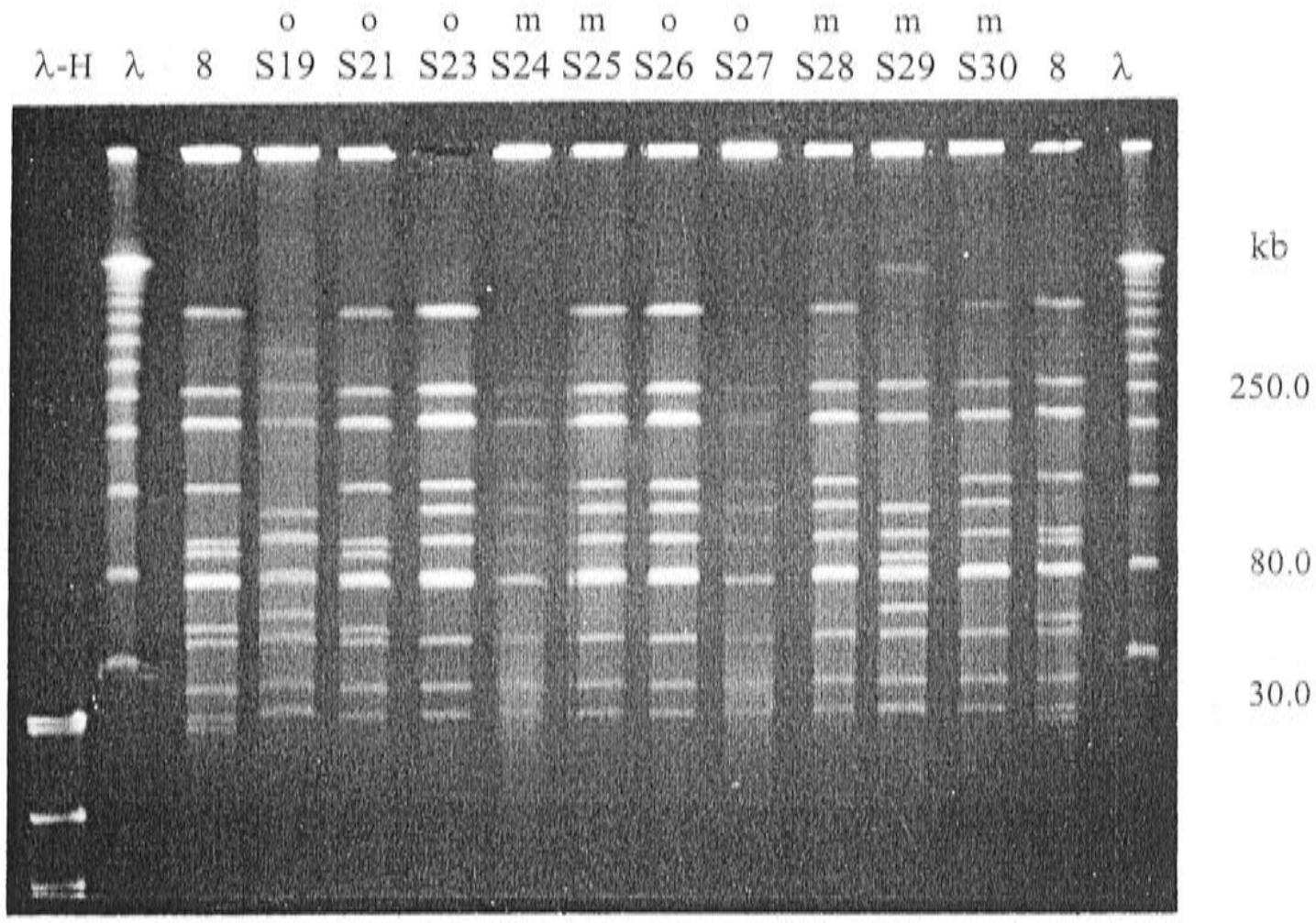


Figure 4.5 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 10 Hib isolates recovered from the cerebrospinal fluid or blood of meningitis (m) patients or patients with other (o) infections from the Sydney region. Lanes labelled 8, λ and λ-H contain fragment standards.

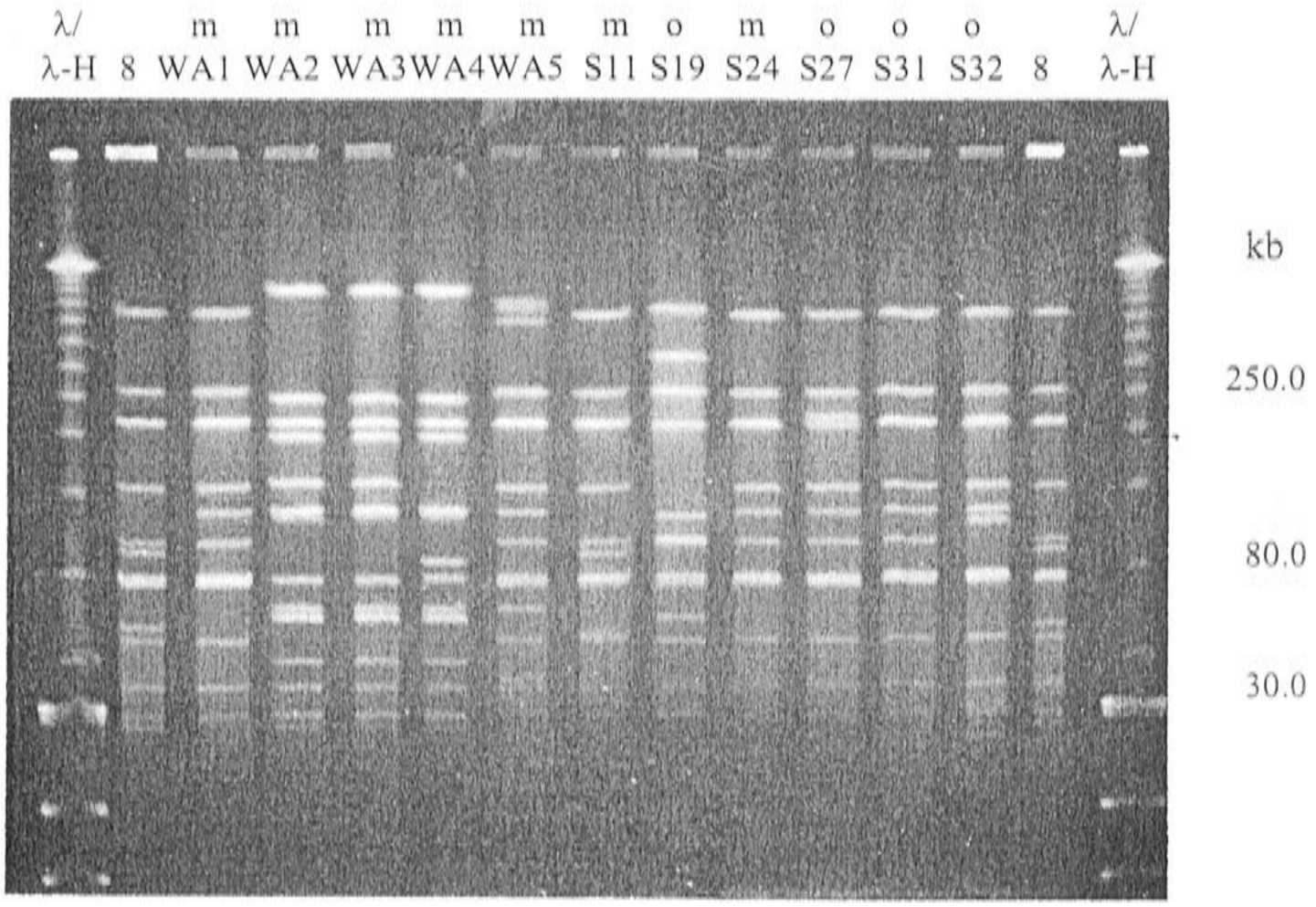


Figure 4.6 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 6 Hib isolates recovered from the cerebrospinal fluid or blood of meningitis (m) patients or patients with other (o) infections from the Sydney region and 5 Aboriginal meningitis isolates from Western Australia. Lanes labelled 8 and λ/λ-H contain fragment standards.

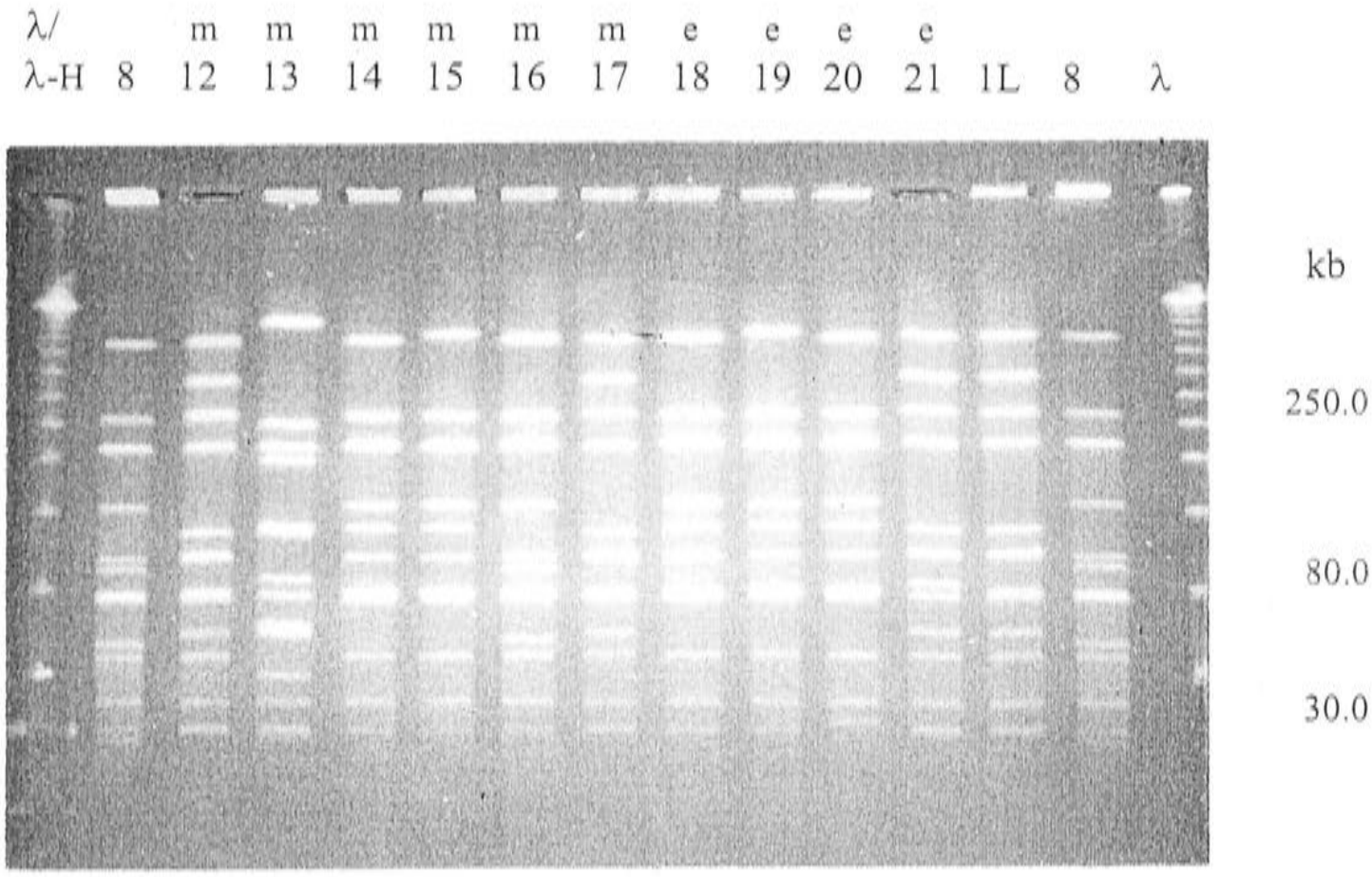


Figure 4.7 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of Hib isolates recovered from the cerebrospinal fluid or blood of 6 patients with meningitis and 4 with epiglottitis from Western Australia. Lanes labelled 8 and λ/λ-H contain fragment standards; lane 1L is a Hib isolate with OMP subtype 1L from a set of OMP standards. The prefix for each isolate is WA.

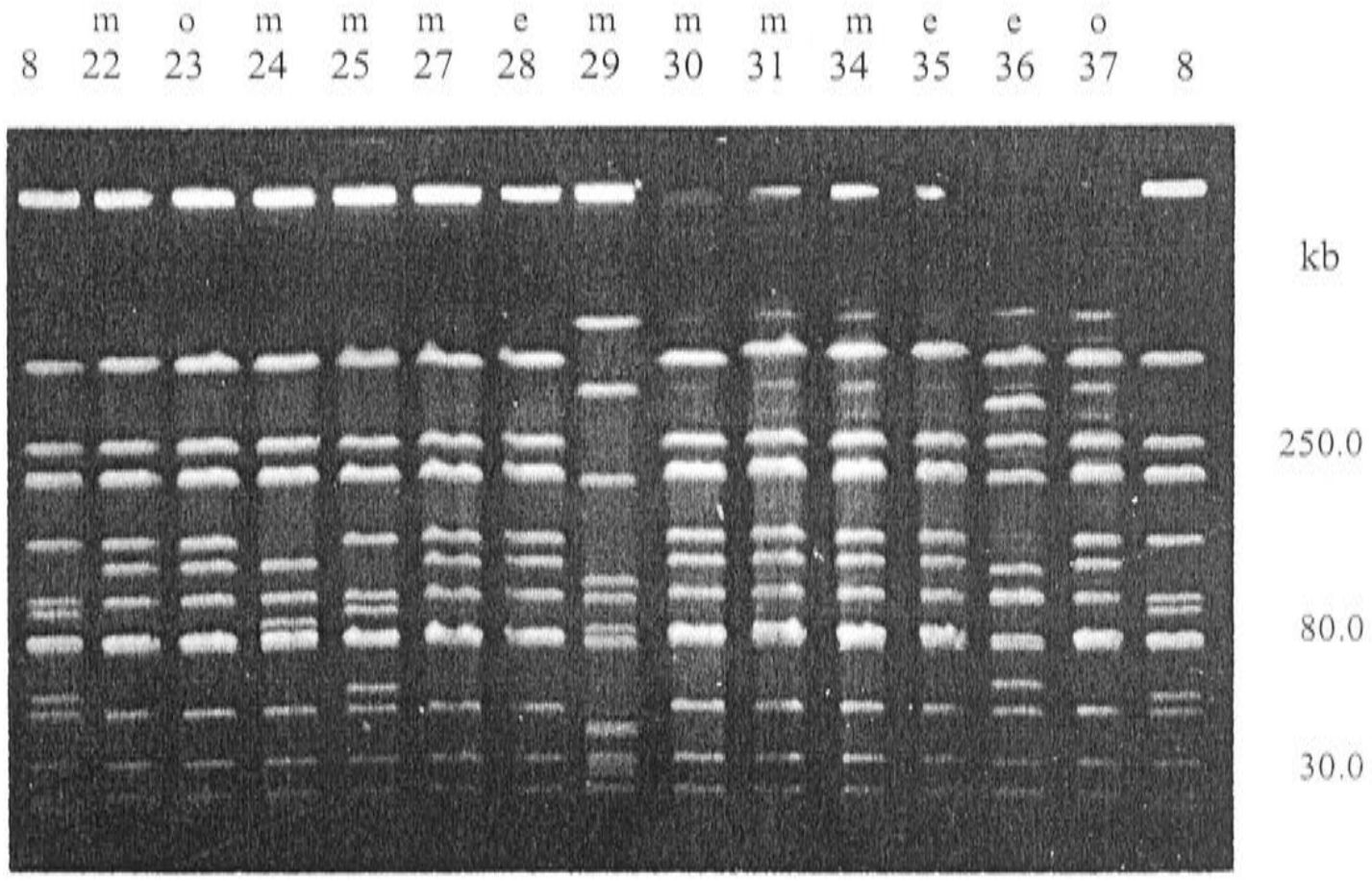


Figure 4.8 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of Hib isolates recovered from the cerebrospinal fluid or blood of 8 meningitis (m) and 3 epiglottitis (e) patients from Western Australia. Two isolates are from patients with other (o) diagnoses. Lanes labelled 8 contain fragment standards. The prefix for each isolate is WA.

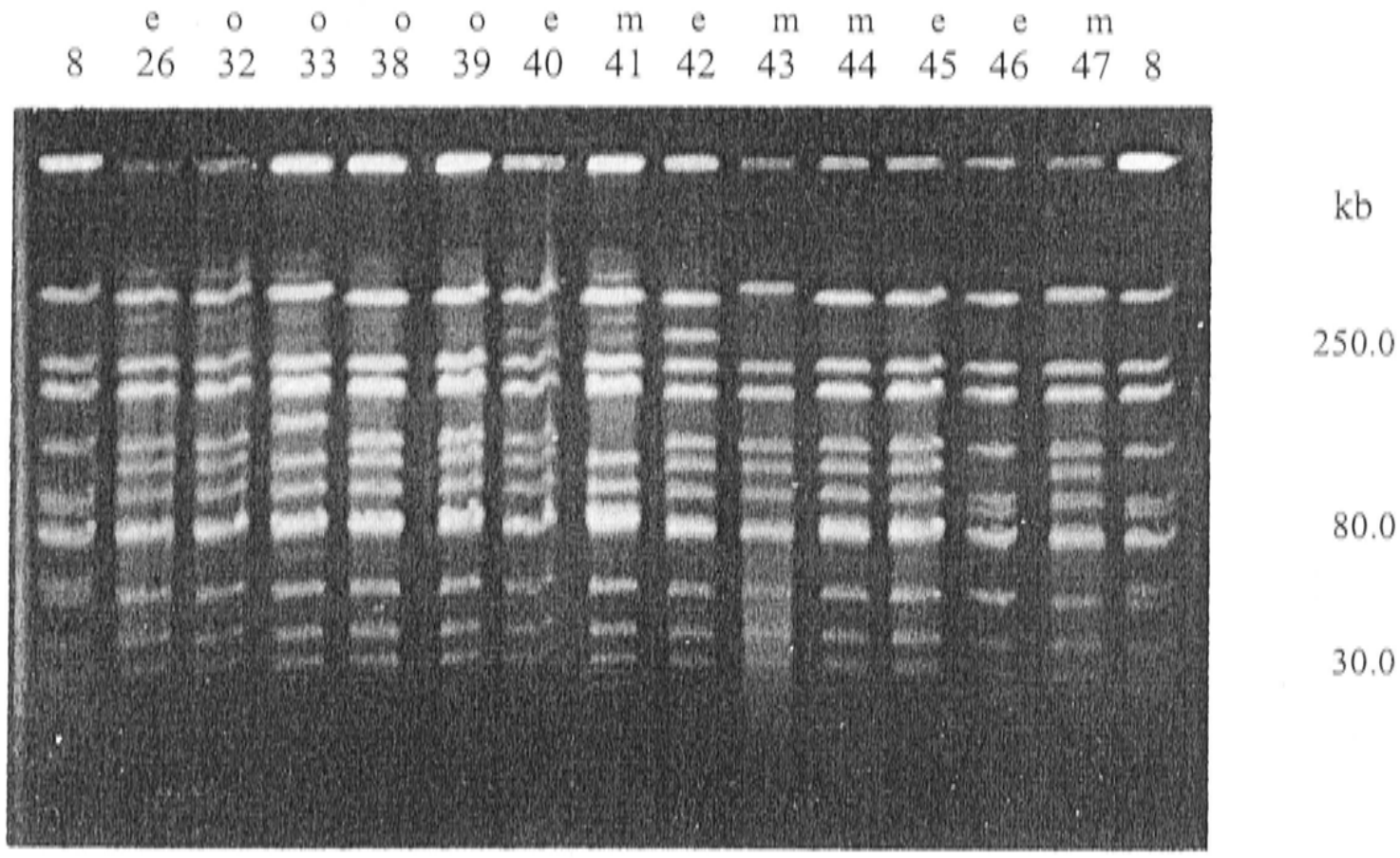


Figure 4.9 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of Hib isolates recovered from the cerebrospinal fluid or blood of 4 meningitis (m) and 5 epiglottitis (e) patients from Western Australia. Four isolates are from patients with other (o) diagnoses. Lanes labelled 8 contain fragment standards. The prefix of all the isolates is WA.

to 6 types (WA15, WA16, WA20, WA24, WA42, and WA46) and together these types comprised 89% of the isolates. Three types (WA14, WA15, and WA17) were shared by both the meningitis and epiglottitis isolates. The lowest ratio of types to isolates of all the urban areas was found in this sample. The distribution of isolates among types was:

Meningitis and epiglottitis isolates combined

*Sma*I-RFLP type	Number of isolates/type	Distribution of isolates/type/disease	
		Meningitis	Epiglottitis
WA14*	12	7	5
WA15	6	4	2
WA17	2	1	1
WA16	2	2	0
WA24	2	2	0
all others (*n=4*)	1 each	—	WA20,21,42,46

*indistinguishable from RFLP type *Sma*I-M5 that predominated in the Canberra sample

All isolates from the four areas (Table 4.5) Thirty-two *Sma*I-RFLP types among the 67 meningitis isolates and 21 types among the 37 epiglottitis isolates were found. Over one third of the sample were members of the predominant type M5. The distribution of isolates among types is shown in Table 4.5.

When all the isolates from all four urban regions were combined 47 types were found. The predominant type, M5 (*n=37*), and its 18 clonally related types (M2, M3, M8, M11, M17, M19, E2, E9, V5, V10, V14, V19, S2, S11, S20, WA24, WA42 and WA46) comprised 67% of the sample. Twelve types had multiple members. Six of them were shared by both the meningitis and the epiglottitis isolates and 6 types had two or three members of only one disease state. But no one type was associated with meningitis or epiglottitis.

The consistency in the ratio of types to isolates found in the different areas or in disease association varied. The ratios among urban areas and disease types was:

Ratio of *Sma*I-RFLP types to isolates

	Meningitis	Epiglottitis	Combined
Canberra	0.80	0.86	0.76
Melbourne	0.60	0.40	0.45
Sydney	0.52	—	0.45
Perth	0.31	0.58	0.32
All urban areas	0.48	0.57	0.45

Canberra had the highest ratio of types to isolates among both meningitis and epiglottitis isolates. Western Australia had the lowest ratio among meningitis isolates

(and overall), while Victoria had the lowest ratio among epiglottitis isolates. In contrast to these observations, the ratio among Victoria's meningitis isolates was twice that found in Western Australia and the ratio of epiglottitis isolates in Western Australia was higher than that in Victoria. Overall the ratio of types to isolates in the entire sample was similar to that found in Victoria and Sydney but was less than that in Canberra and slightly more than in Western Australia. It is interesting that the two subsets that had the largest difference in ratios, Canberra and Western Australia, had the largest sample sizes. If the samples examined from Victoria and Sydney are smaller proportions of the total number of isolates in those areas than are the Canberra and Western Australian samples they may be less representative of the true diversity of Hib. This may explain the low ratios seen in these two areas.

A breakdown of the distribution of types, multiple isolate types, and shared types in the entire sample is shown in Table 4.5. Only one type, *Sma*I-M5 was found in all four urban areas. Six other types were found in more than one area.

4.3.3 Genetic relationships determined by numerical analysis of restriction fragment length polymorphisms

The range of similarity or genetic distance, identification of clonal relationships and identification of the most diverse relationships among types were determined from matrices produced using the mathematical model of Nei and Li except for the subset of epiglottitis isolates from Western Australia. In this case, Jaccard's algorithm was used because the TDraw program would not compute the Nei and Li matrix.

Canberra isolates (Figures 3.9-3.11) The genetic analysis of these isolates is described in Chapter 3. Briefly, a wide range of similarity among the types was found in this sample. Three types were genetically distant from each other and the major clusters into which most of the types fell. Two of these, M13-M7, had an *F* value of 0.19. A predominant type, designated *Sma*I-M5, and 6 clonally related types accounted for 55% of the sample.

Victorian isolates (Figures 4.10-4.12) The similarity among pairs of meningitis types ranged from 50 to 92%. The predominant type and 2 clonally related types represented 60% of this subset. The most genetically distinct type had *F* values ranging from 0.50 to 0.57 when compared to the other types. The epiglottitis types ranged in similarity from

Table 4. 5 Summary of *Sma*I-RFLP types[l] found among 104 non-Aboriginal meningitis and epiglottitis isolates of *Haemophilus influenzae* type b

RFLPs among meningitis isolates *n=67 isolates*		RFLPs among epiglottitis isolates *n=37 isolates*		RFLPs among combined meningitis and epiglottitis isolates *n=104 isolates*		
M1	**V1**[d]	E1	E13	M1	E1	V18
M2	V5	**E2**[i]	E15	M2	**E2***[n]	V19
M3*[a]	V7	E3	V14	**M3***[a]	E3	S2
M5*[b]	V8	E4	V18	**M5***[b,k]	E4	S5
M6	V10	E5	V19	M6	E5	S11
M7	S2	E6	S37	M7	E6	S12
M8	S5	E7	**WA19**[l]	**M8**[m]	**E9***[j]	**S16**[f]
M9	**S7**[e]	**E9***[j]	WA21	M9	**E11***[o]	**S20***[g,l]
M10	S8	E11	WA36	M10	E13	S29
M11	S11	**E12***[k]	WA42	M11	E15	S37
M13	S12		WA46	M13	**V1**[d]	**WA17**[p]
M14	**S16**[f]		*n=21*	M14	V5	WA21
M15*[c]	**S20***[g]			**M15***[c]	V7	**WA24**[h]
M17	S29			M17	V8	WA42
M19	WA17			M19	V10	WA46
M20	**WA24**[h]			M20	V14	*n=47*
	n=32					

[a]M3, M4, WA16, WA25
[b]M5,M12,M16,M18,V2, V3,V4,V6,S6,S17, S18, S24,S25,S28,S30,S35, S36,WA14,WA22, WA27,WA29,WA30, WA44,WA47 *(n=24)*
[c]M15, S14
[d]V1, V9
[e]S7, S9
[f]S16, S22
[g]S20, WA15, WA31, WA34, WA43
[h]WA24, WA41
[i]E2, E14
[j]E9, E10,WA20
[k]E12,V11,V12,V13,V15, V16,V17,V20,WA18, WA26,WA28,WA40, WA45 *(n=13)*
[l]WA19, WA35
[m]M8, E7
[n]E2, E14, S8
[o]E11, S7, S9
[p]WA17, WA36

[l]The designated RFLP type name represents the lowest number assigned to a meningitis isolate of that type or if no meningitis isolate represents the type then the lowest number assigned to an epiglottitis isolate of that type. Type names were assigned to Canberra isolates first, then new types found among Melbourne isolates, then Sydney, and finally Perth. The types in boldface represent those types with multiple isolates. The boxed types are the types found among both the meningitis and the epiglottitis isolates.
*Found in more than one urban area; RFLP-M5 is the only type found in all four urban areas.

79 to 93%. Three of the 4 types in this subset were clonally related and accounted for 90% of the isolates. It is possible that the most diverse isolate, V18 (F=0.79 when compared with V11 and V14 and 0.80 when paired with V19), was clonally related to the others. When the meningitis and epiglottitis isolates were combined the predominant type and its clonally related types comprised 75% of the sample. The most genetically

distinct type was no greater than 50% dissimilar from other types and overall this sample was more homogeneous than the Canberra sample.

Sydney isolates (Figure 4.13) The meningitis types ranged in similarity from 55 to 96%. The predominant type and 4 clonally related types accounted for 62% of the isolates. Another clonal group of 3 types comprising 4 members was 70% similar to the predominant type.

The epiglottitis isolate was the most diverse in the combined sample and the largest *F* value, 0.27, was found when it was compared to S11. This reduced the similarity range to a low of 27% for the combined sample. The large genetic distance between some types was similar to that found in the Canberra sample.

Western Australian isolates (Figures 4.14-4.16) The similarity range of the meningitis types was 60 to 93%. The predominant type and clonally related types ($n=4$) accounted for 93% of the sample. Among the epiglottitis types a larger similarity range of 46 to 96% was found. Clonally related types including the predominant type represented 75% of the epiglottitis isolates. When combined the predominant type and its clonally related types comprised 89% of the sample.

All isolates from the four urban areas (Figures 4.17-4.19) The range of similarity among all the meningitis types was determined by the diversity of the Canberra types and ranged from 19 to 97%. The predominant type and its clonally related types comprised 45 members and accounted for 67% of the meningitis isolates. Four meningitis types were less the 50% similar to the other 28 types. M13-M7 was the most diverse pair with an $F=0.19$.

Most of the epiglottitis isolates fell into 2 closely related clusters. The predominant type and its clonally related types accounting for 68% of the sample represented twenty-five isolates. S37, the most diverse type, was less then 40% similar to all the other types. S37-E4 was the most diverse pair of epiglottitis types with an $F=0.20$.

When all the types were combined they fell into 7 major clusters in a tree like dendrogam. Three clusters fell at an $F\leq 0.50$ and 25 branches fell at $F\geq 0.90$. Six types, M13, S37, M7, V1, M9 and WA21 represented the most diverse types and included 7

isolates. The predominant type, designated *Sma*I-M5 (*n=37*) and its 18 clonally related types accounted for 70 isolates or 67% of the entire sample.

The distribution of RFLP type *Sma*I-M5 and its clonally related types among the areas and subsets was:

Percentage of sample represented by *Sma*I-M5 and clonally related types*

	Meningitis	Epiglottitis	All isolates
Canberra	20/55	7/43	15/47
Victoria	40/60	70/90	55/75
Sydney	43/62	—	41/59
Western Australia	44/94	42/50	43/89
All areas	36/67	35/68	35/67

*the first number of a pair refers to the percentage of the predominant type alone; the second number refers to the percentage of the predominant type and its clonally related types.

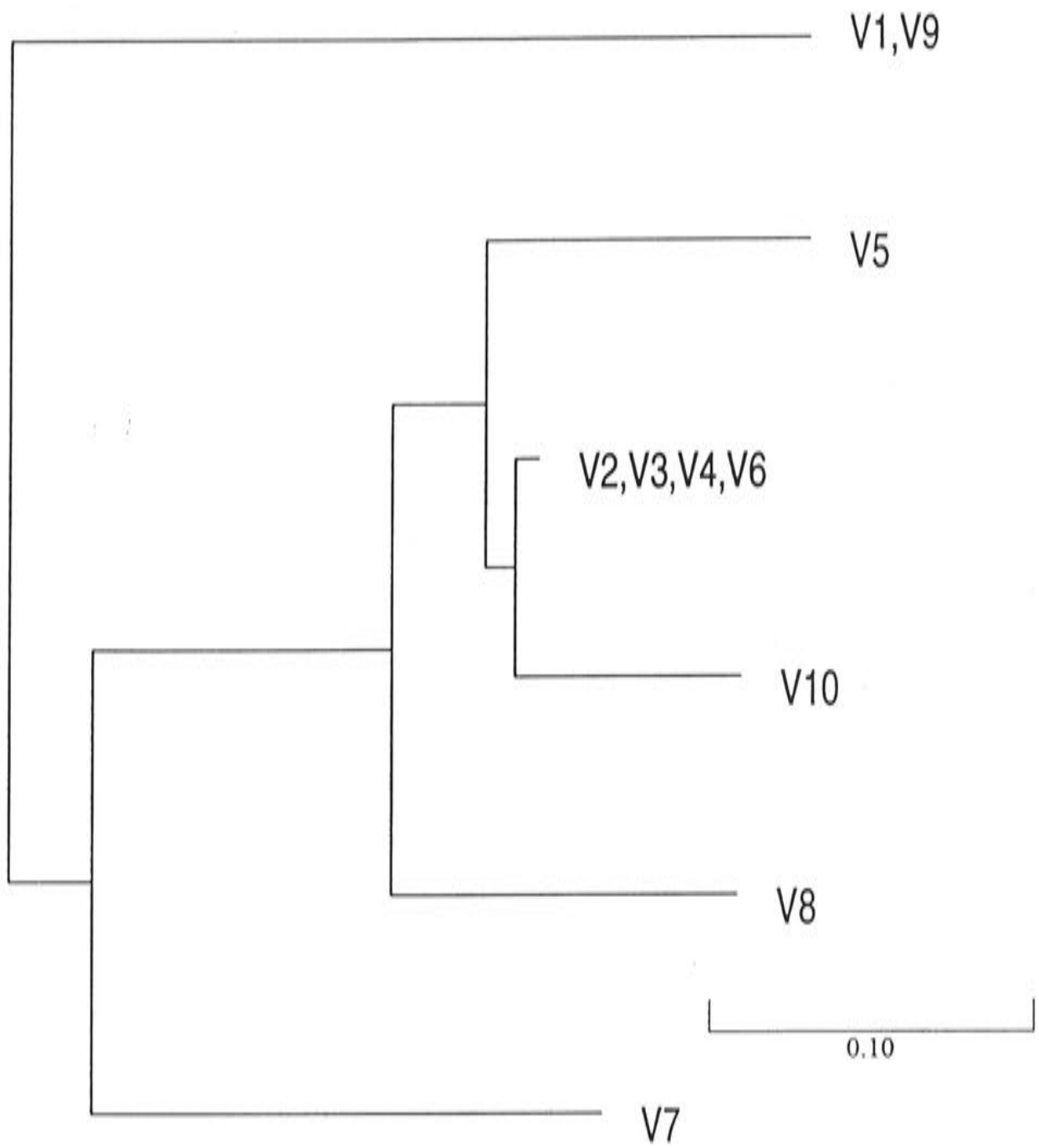


Figure 4.10 **Dendrogram showing the clustering of 6 *Sma*I-RFLP types found among 10 meningitis isolates obtained from non-Aboriginal patients in Victoria.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the horizontal lines between pairs is an estimate of the genetic distance between the two types.

Types that fell in clusters ($n=2$) at $F \leq 0.5$	Types that fell in the clusters or branches ($n=5$) at $F \geq 0.9$	Pairs of types ($n=2$) at $F \geq 0.9$
1) V1,V9	1) V2[a],V10	V2-V5
2) All the other types	2) V5	V2-V10
	3) V8	
	4) V7	
	5) V1,V9	

[a]This type was indistinguishable from *Sma*I-M5, the predominant type found among the Canberra isolates.

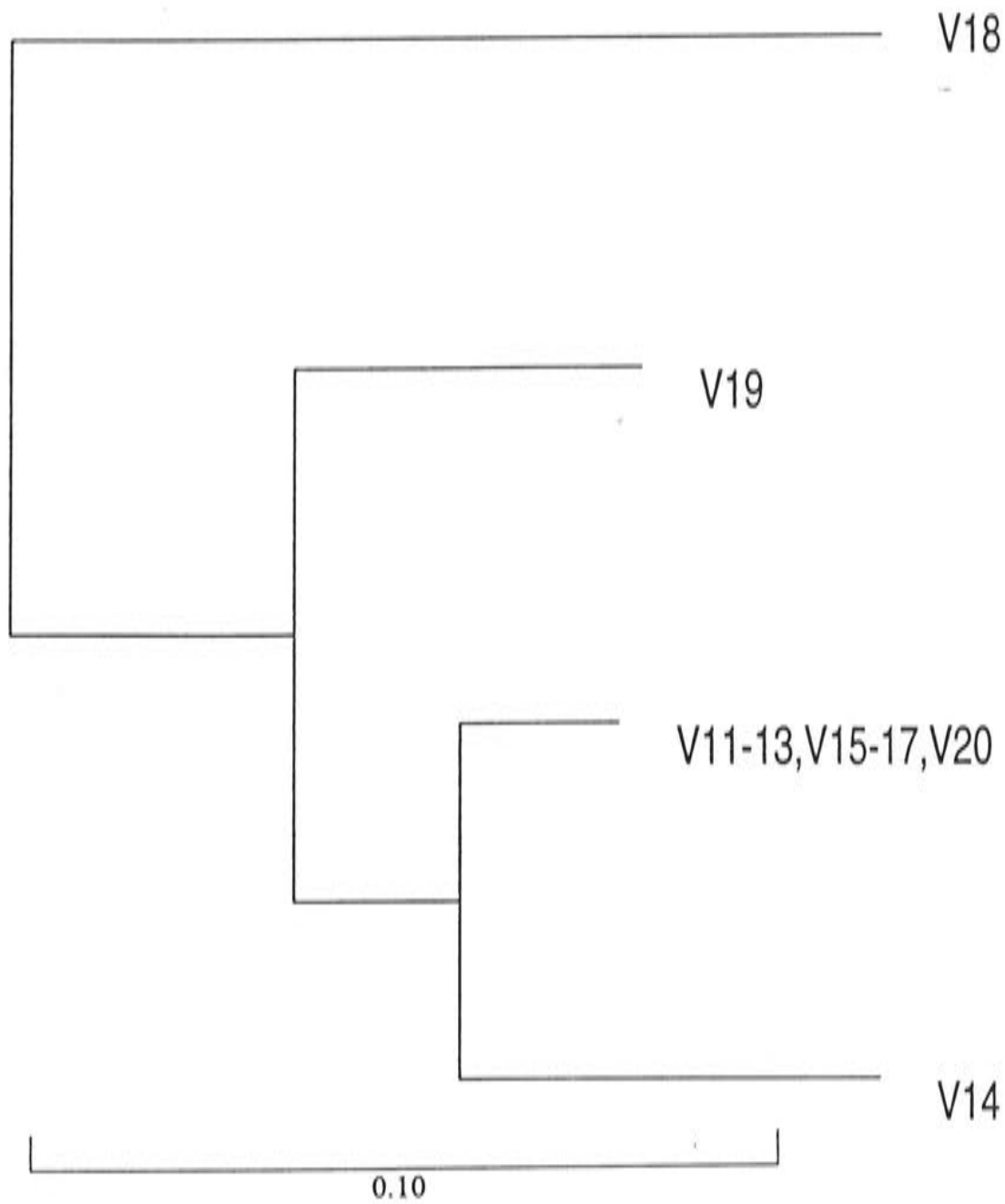


Figure 4.11 **Dendrogram showing the clustering of 4 *Sma*I-RFLP types found among 10 epiglottitis isolates obtained from non-Aboriginal patients in Victoria.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the horizontal lines between pairs is an estimate of the genetic distance between the two types.

Types that fell in clusters ($n=1$) at $F \leq 0.5$	Types that fell in the clusters or branches ($n=3$) at $F \geq 0.9$	Pairs of types ($n=2$) at $F \geq 0.9$
1) All the types	1) V11[a],V19	V11-V19
	2) V14	V11-V14
	3) V18	

[a]This type was indistinguishable from *Sma*I-M5, the predominant type found among the Canberra isolates.

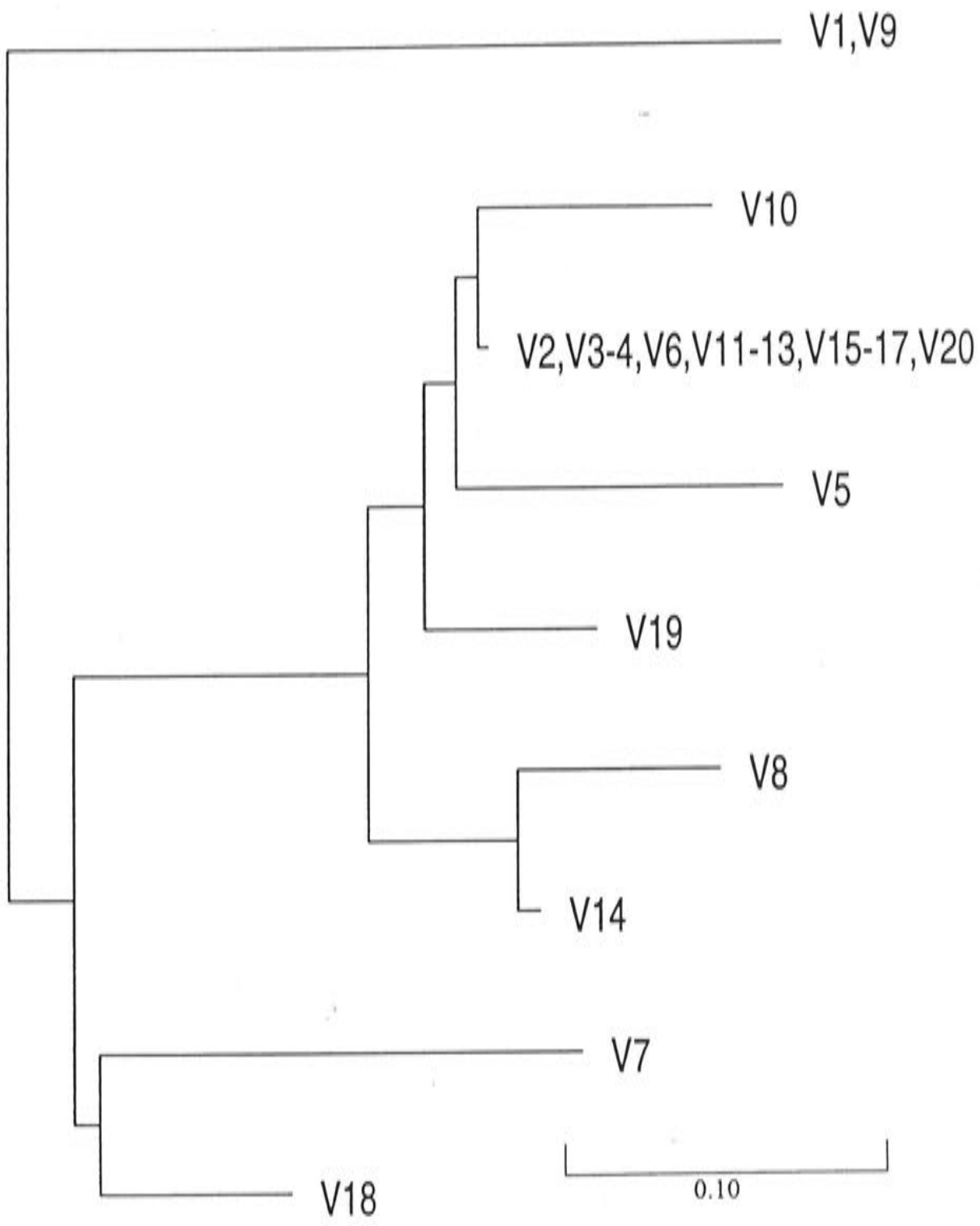


Figure 4.12 **Dendrogram showing the clustering of 9 *Sma*I-RFLP types found among 20 meningitis and epiglottitis isolates obtained from non-Aboriginal patients in Victoria.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the horizontal lines between pairs is an estimate of the genetic distance between the two types.

Types that fell in clusters ($n=2$) at $F \leq 0.5$	Types that fell in the clusters or branches ($n=7$) at $F \geq 0.9$	Pairs of types ($n=5$) at $F \geq 0.9$
1) V1,V9	1) V2[a],V19	V2-V5
2) All the other types	2) V5	V2-V10
	3) V10	V2-V14
	4) V8,V14	V2-V19
	5) V7	V8-V14
	6) V18	
	7) V1,V9	

[a]This type was indistinguishable from *Sma*I-M5, the predominant type found among the Canberra isolates.

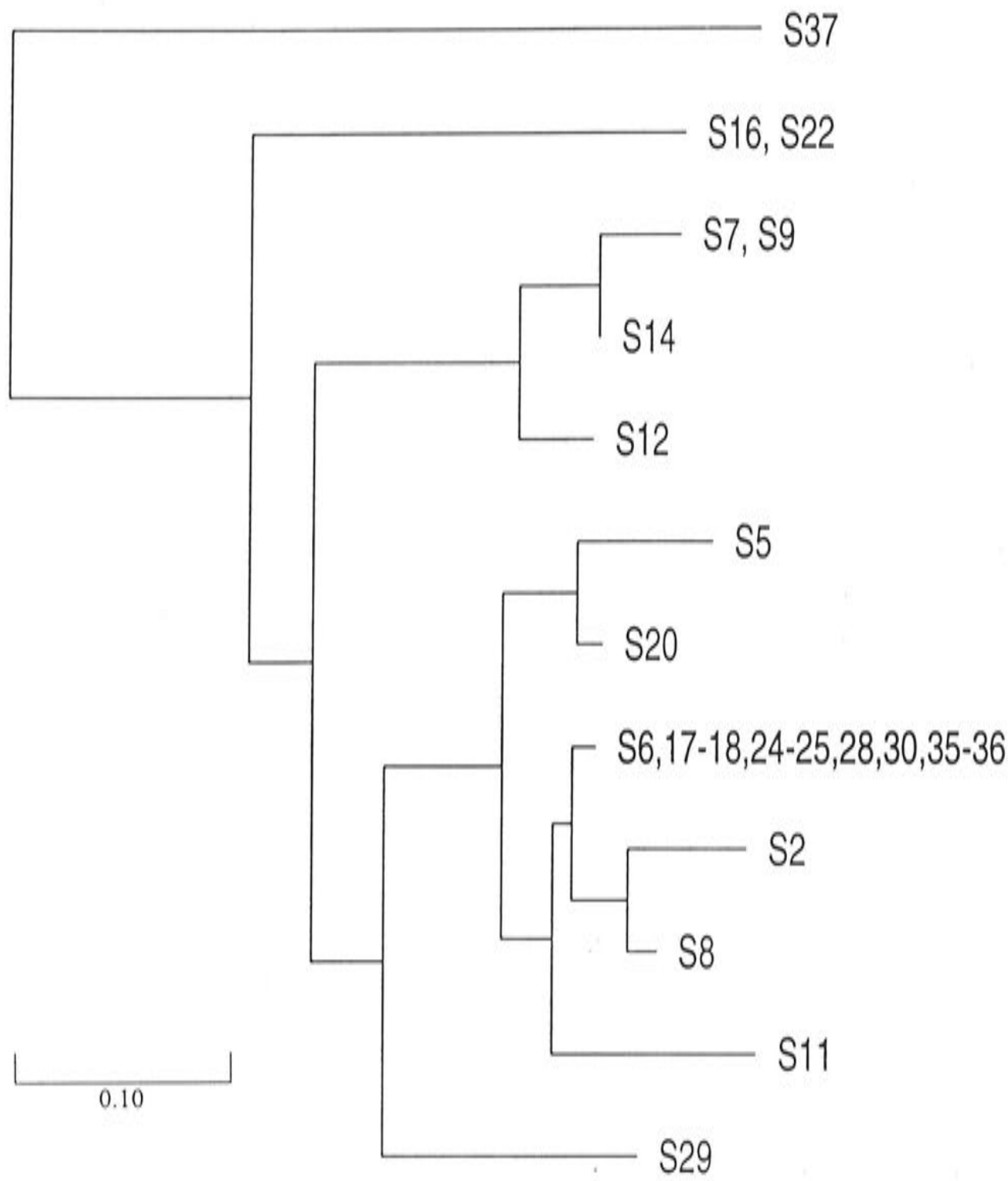


Figure 4.13 **Dendrogram showing the clustering of 12 *Sma*I-RFLP types found among 22 meningitis and epiglottitis isolates obtained from non-Aboriginal patients in the Sydney region.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the horizontal lines between pairs is an estimate of the genetic distance between the two types.

Types that fell in clusters ($n=2$) at $F \leq 0.5$	Types that fell in the clusters or branches ($n=7$) at $F \geq 0.9$	Pairs of types ($n=11$) at $F \geq 0.9$	
1) S37	1) S6[a],S2,S8	S2-S6	S6-S20
2) All the other types	2) S11	S2-S8	S7-S12
	3) S5,S20	S2-S20	S7-S14
	4) S29	S5-S20	S8-S20
	5) S7,S14,S12	S6-S8	S12-S14
	6) S16	S6-S4	
	7) S37		

[a]This type was indistinguishable from *Sma*I-M5, the predominant type found among the Canberra isolates.

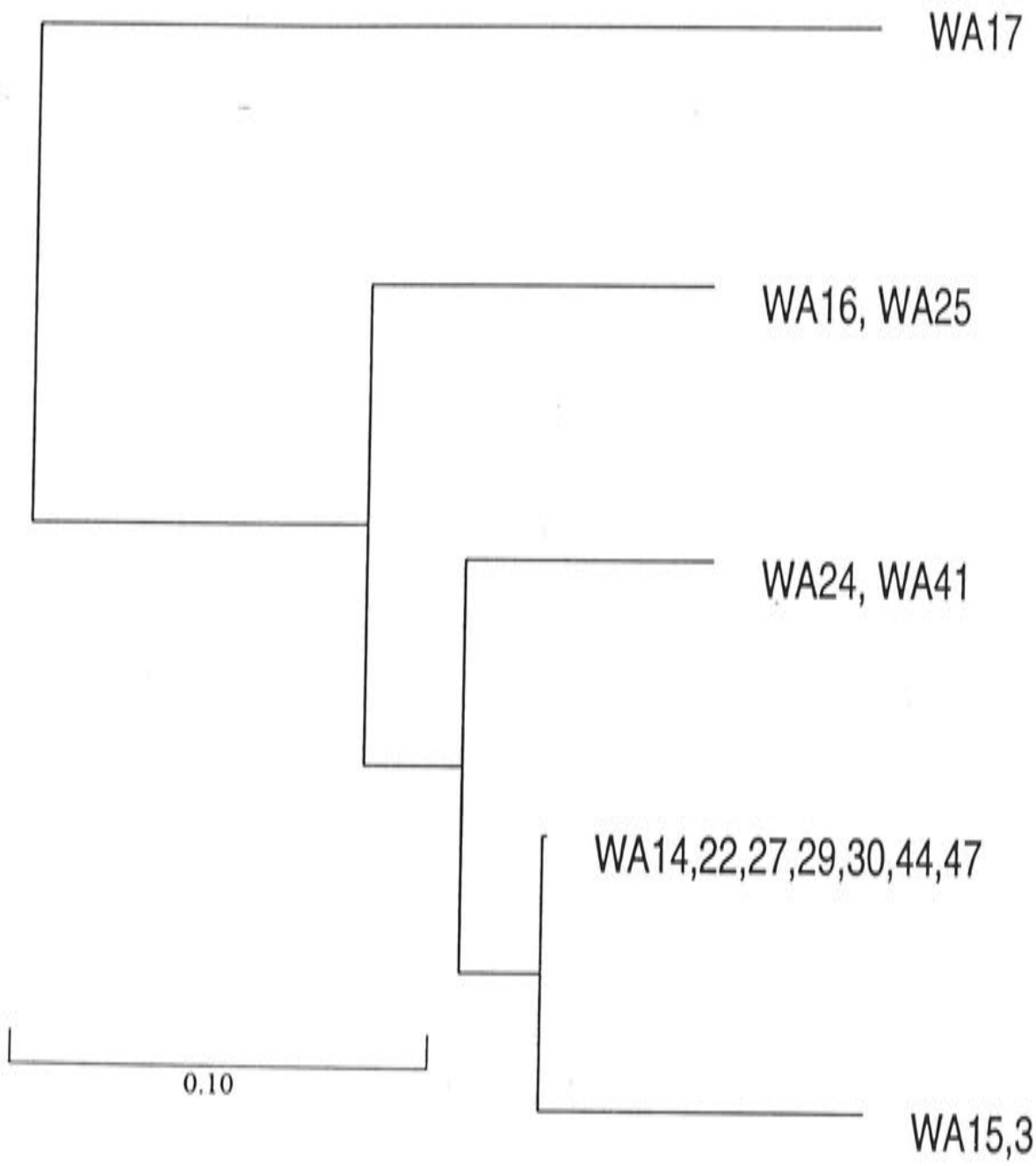


Figure 4.14 **Dendrogram showing the clustering of 5 *Sma*I-RFLP types found among 16 meningitis isolates obtained from non-Aboriginal patients in Western Australia.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the horizontal lines between pairs is an estimate of the genetic distance between the two types.

Types that fell in clusters ($n=1$) at $F \leq 0.5$	Types that fell in the clusters or branches ($n=4$) at $F \geq 0.9$	Pairs of types ($n=3$) at $F \geq 0.9$
1) All the types	1) WA14[a],WA15	WA14-WA15
	2) WA24	WA14-WA16
	3) WA16	WA14-WA24
	4) WA17	

[a]This type was indistinguishable from *Sma*I-M5, the predominant type found among the Canberra isolates.

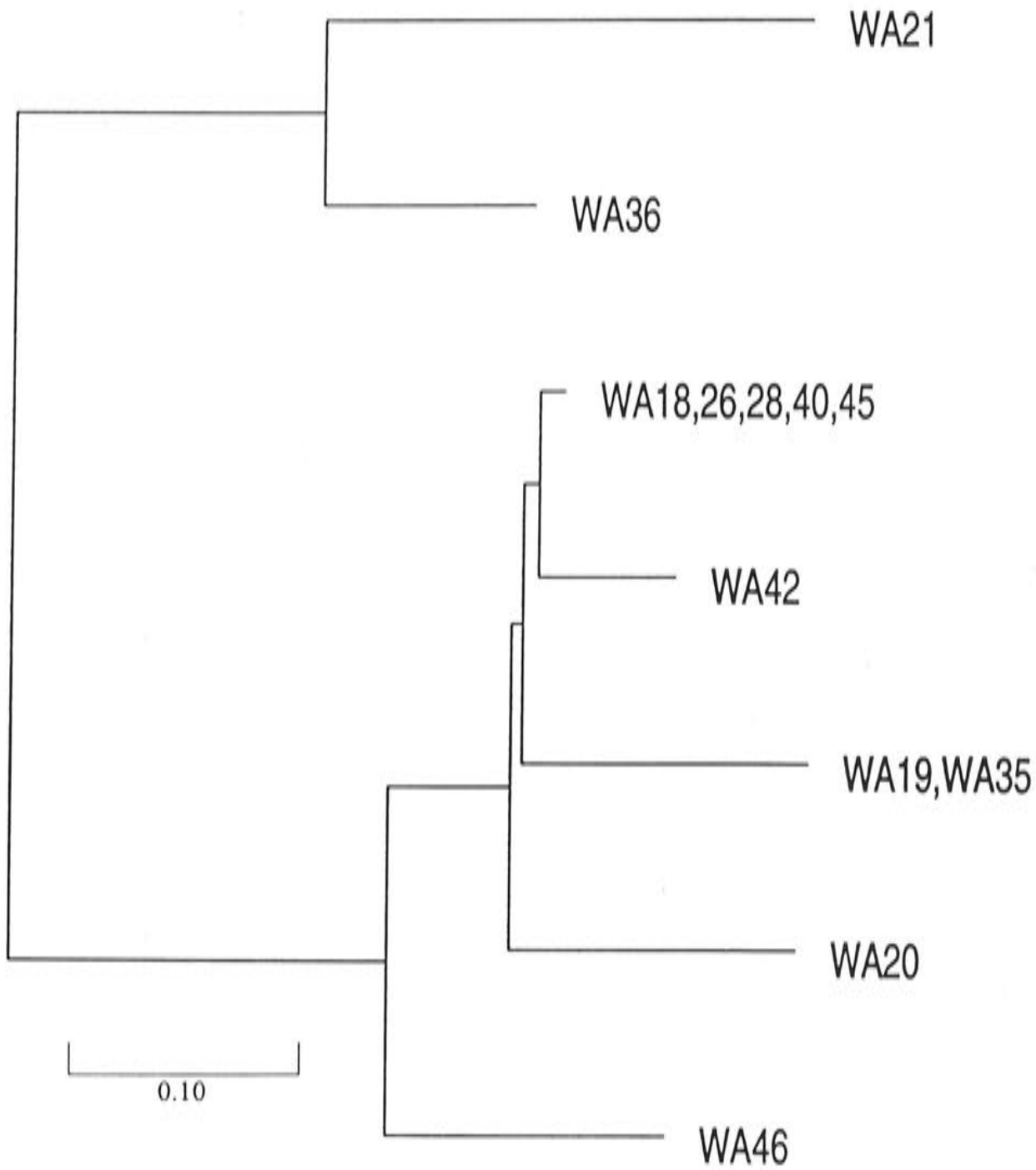


Figure 4.15 **Dendrogram showing the clustering of 7 *Sma*I-RFLP types found among 12 epiglottitis isolates obtained from non-Aboriginal patients in Western Australia.** The dendrogram was generated using Jaccard's mathematical model. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the horizontal lines between pairs is an estimate of the genetic distance between the two types.

Types that fell in clusters ($n=2$) at $F \leq 0.5$	Types that fell in the clusters or branches ($n=6$) at $F \geq 0.9$	Pairs of types ($n=1$) at $F \geq 0.9$
1) WA21,WA36	1) WA18[a],WA42	WA18-WA42
2) All the other types	2) WA19	
	3) WA20	
	4) WA46	
	5) WA36	
	6) WA21	

[a]This type was indistinguishable from *Sma*I-M5, the predominant type found among the Canberra isolates.

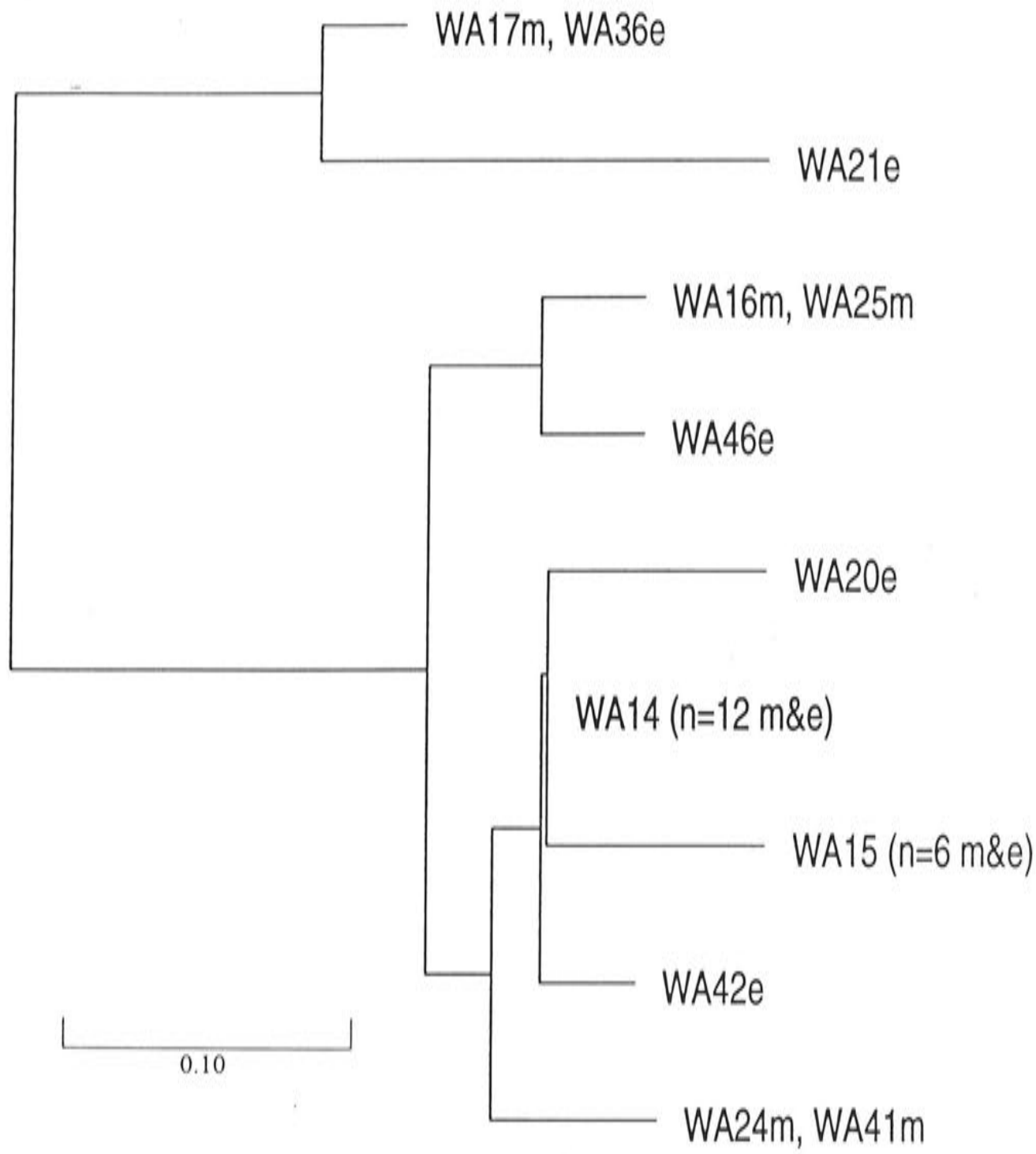


Figure 4.16 **Dendrogram showing the clustering of 9 *Sma*I-RFLP types found among 28 meningitis and epiglottitis isolates obtained from non-Aboriginal patients in Western Australia.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the horizontal lines between pairs is an estimate of the genetic distance between the two types. An 'm' for meningitis or 'e' for epiglottitis is used to indicate the disease association of an isolate.

Types that fell in clusters ($n=2$) at $F \leq 0.5$	Types that fell in the clusters or branches ($n=6$) at $F \geq 0.9$	Pairs of types ($n=10$) at $F \geq 0.9$	
1) WA21	1) WA14[a],WA15[b],WA42	WA14-WA15	WA14-WA46
2) All the other types	2) WA20	WA14-WA16	WA15-WA42
	3) WA24	WA14-WA20	WA16-WA46
	4) WA16,WA46	WA14-WA24	WA20-WA42
	5) WA17	WA14-WA42	WA24-WA42
	6) WA21		

[a]This *Sma*I-RFLP type included isolates WA14m, WA22m, WA27m, WA29m, WA30m, WA44m, WA47m, WA18e, WA26e, WA28e, WA40e and WA45e ($n=12$) and was indistinguishable from *Sma*I-M5, the predominant type found among the Canberra isolates.

[b]This *Sma*I-RFLP type included isolates WA15m, WA31m, WA34m, WA43m, WA19e and WA35e ($n=6$).

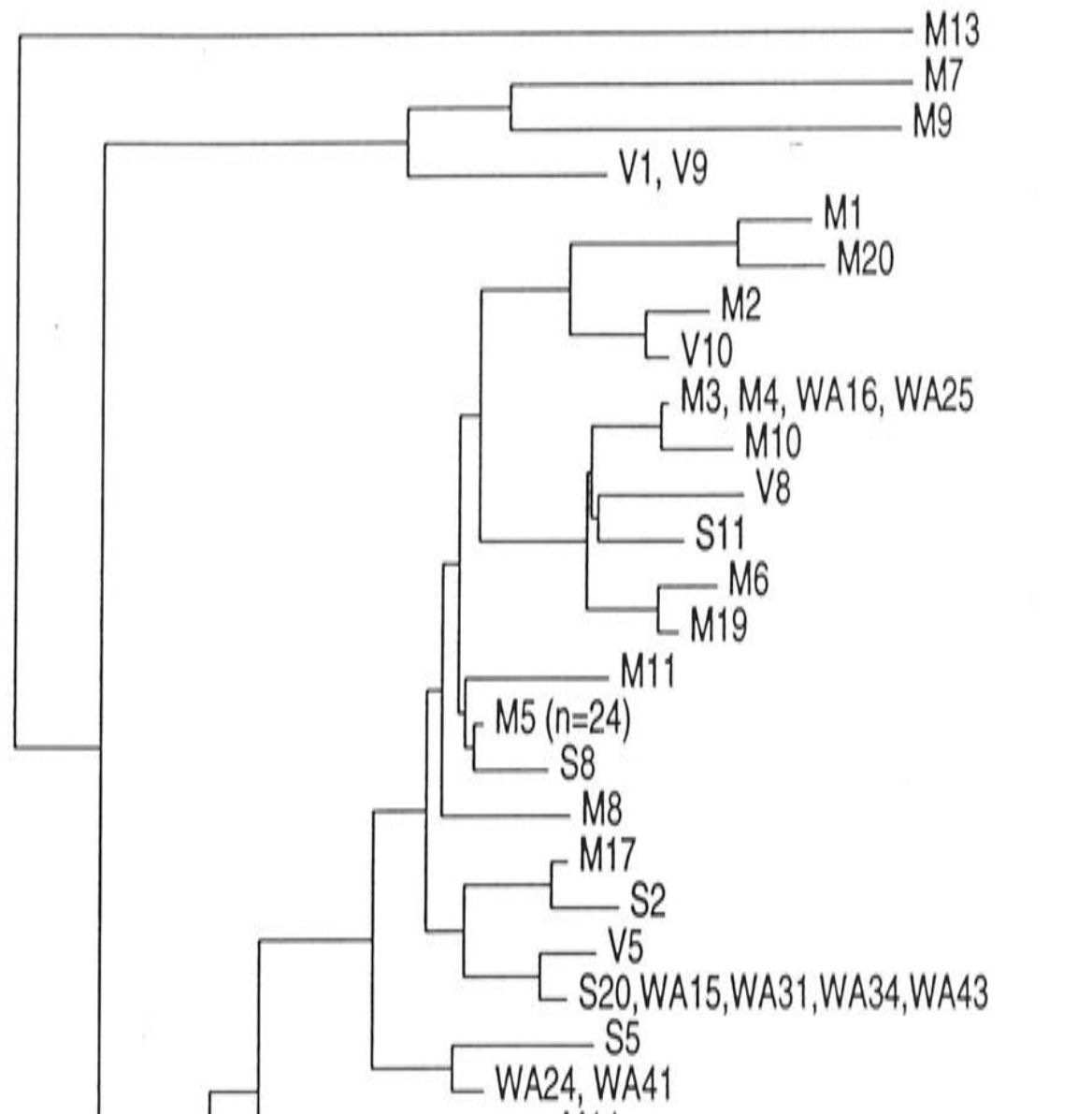


Figure 4.17 **Dendrogram showing the clustering of 32 *Sma*I-RFLP types found among 67 meningitis isolates obtained from non-Aboriginal patients in four urban regions.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10.

Types that fell in clusters ($n=3$) at $F \leq 0.5$

1) M13
2) M7, M9 and V1
3) All the other types

Types that fell in the clusters or branches ($n=17$) at $F \geq 0.9$

1) M1,M20
2) M2,V10
3) M3,M10
4) V8,S11
5) M6,M19
6) M5[a],M11,S8
7) M8
8) S5,WA24
9) M17,S2,V5,S20
10) M14
11) S16
12) M15,S7,WA17
13) S12
14) S29
15) V7
16) M7,M9,V1
17) M13

Pairs of types ($n=49$) at $F \geq 0.9$

M1-M20	M5-M17	M10-S11	V8-S11
M1-V10	M5-M19	M11-S8	V10-S8
M2-M5	M5-V5	M15-S7	S2-S8
M2-V10	M5-V10	M15-S12	S2-S20
M3-M5	M5-S2	M15-WA17	S5-S20
M3-M6	M5-S8	M17-V5	S5-WA24
M3-M10	M5-S11	M17-S20	S7-S12
M3-M11	M5-S20	M17-S2	S7-WA17
M3-M19	M5-WA24	M17-S8	S8-S20
M3-V8	M6-M19	M19-V8	S8-WA24
M3-S11	M6-S11	M19-S11	
M5-M8	M8-V5	V5-S5	
M5-M11	M8-S8	V5-S20	

[a]This type included isolates M5, M12, M6, M18, V2, V3, V4, V6, S6, S17, S18, S24, S25, S28, S30, S35, S36, WA14, WA22, WA27, WA29, WA30, WA44, WA47 ($n=24$).

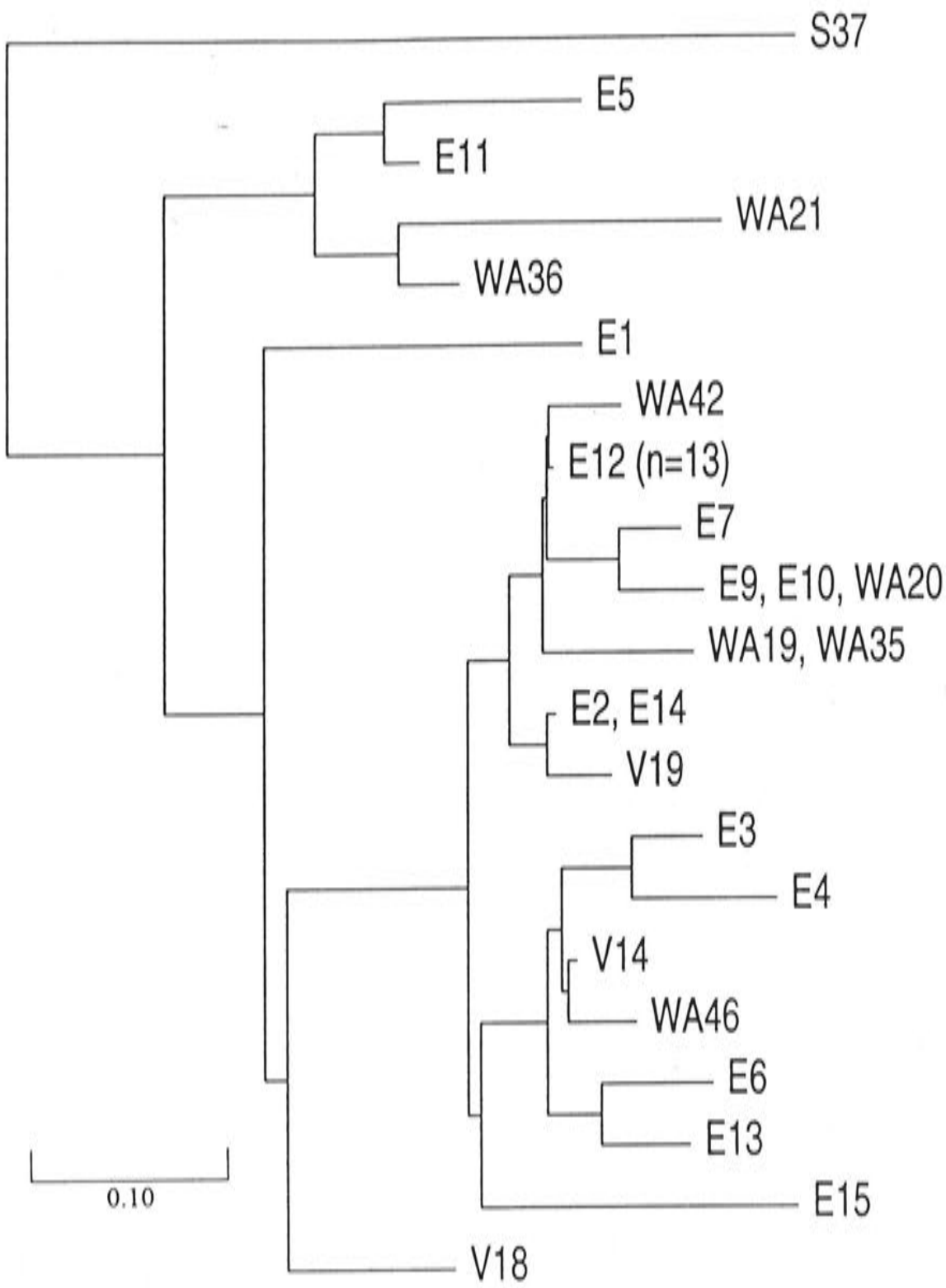


Figure 4.18 **Dendrogram showing the clustering of 21 *Sma*I-RFLP types found among 37 epiglottitis isolates obtained from non-Aboriginal patients in four urban regions.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10.

Types that fell in clusters (*n=2)* at $F \leq 0.5$	Types that fell in the clusters or branches (*n=11*) at $F \geq 0.9$		Pairs of types (*n=31)* at $F \geq 0.9$			
1) S37	1) WA42,E12[a],E7,E9	9) WA21	E2-E7	E3-V14	E7-WA42	E12-WA46
2) All the other types	2) WA19	10) WA36	E2-E9	E3-WA46	E9-E12	E13-V14
	3) E2,V19	11) S37	E2-E12	E4-V14	E9-WA42	E13-WA46
	4) E3,E4		E2-V14	E5-E11	E11-WA36	V14-WA42
	5) V14,WA46,E6,E13		E2-V19	E6-E13	E12-V14	V14-WA46
	6) E15		E2-WA19	E6-V14	E12-V19	V19-WA42
	7) V18		E2-WA42	E7-E9	E12-WA19	WA19-42
	8) E5,E11		E3-E4	E7-E12	E12-WA42	

[a]This type included isolates E12, V11, V12, V13, V15, V16, V17, V20, WA18, WA26, WA28, WA40 and WA45 (*n=13*) and was indistinguishable from *Sma*I-M5, the predominant type among the Canberra isolates.

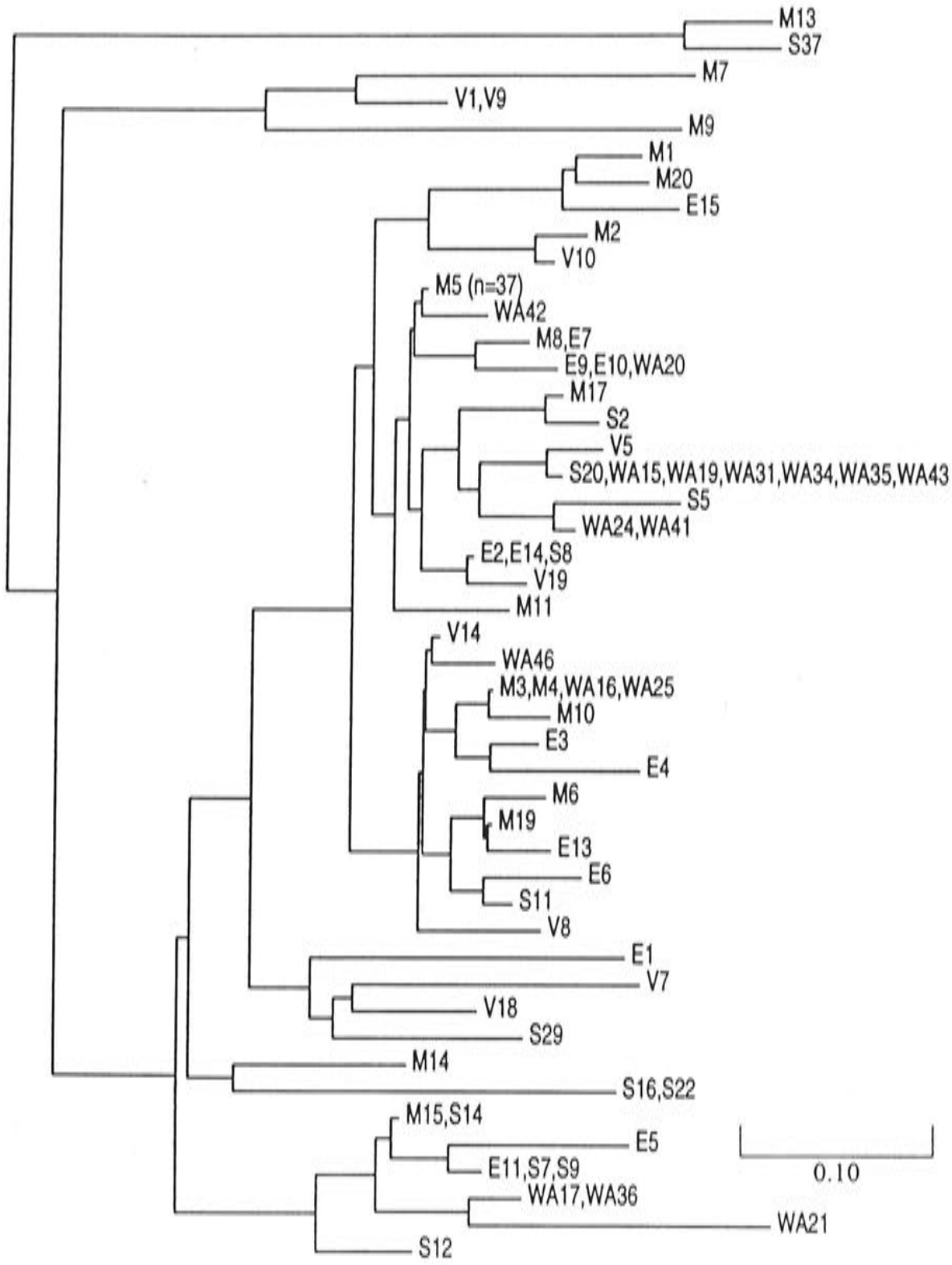


Figure 4.19 **Dendrogram showing the clustering of 47 *Sma*I-RFLP types found among 104 meningitis and epiglottitis isolates obtained from non-Aboriginal patients in four urban regions.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the tree. The bar indicates a genetic value of 0.10.

Types that fell in clusters (*n=3)* at $F \leq 0.5$	Types that fell in the clusters or branches (*n=25*) at $F \geq 0.9$		Pairs of types (*n=104)* at $F \geq 0.9$
1) M13,S37	1) M1,M20,E15	14) E1	Not listed
2) M7,V1, M9	2) M2,V10	15) V7	
3) All the other types	3) M5[a],WA42,M8,E9	16) V18	
	4) M7,S2	17) S29	
	5) V5,S20,S5,WA24	18) M14	
	6) E2,V19	19) S16	
	7) M11	20) M15,E5,E11	
	8) V14,WA46	21) WA17	
	9) M3,M10	22) WA21	
	10) E3,E4	23) S12	
	11) M6,M19,E13	24) M7,V1,M9	
	12) E6,S11	25) M13,S37	
	13) V8		

[a]This type included isolates M5, M12, M6, M18, E12, V2, V3, V4, V6, V11, V12, V13, V15, V16, V17, V20, S6, S17, S18, S24, S25, S28, S30, S35, S36, WA14, WA18, WA22, WA26, WA27, WA28, WA29, WA30, WA40, WA44, WA45, WA47 (*n=37*).

4.3.4 Numerical index of the discriminatory ability of the typing system

Simpson's index of diversity (D) was 0.87 when applied to the data to determine the discriminatory power of *Sma*I among the 104 meningitis and epiglottitis isolates.

4.4 Discussion

A large sample (*n=104*) of meningitis and epiglottitis isolates from four distinct urban regions in Australia was obtained for characterisation by PFGE and hybridisation. It was striking that despite significant heterogeneity among some types, a single type found in all four urban areas and 18 types clonally related to it accounted for well over two-thirds of the isolates. It appears from our results that a group of closely related types accounted for most of the invasive Hib disease in non-Aboriginal Australians in the pre-vaccine era.

These results correlate with findings from Victoria and other areas of the world where outer membrane protein subtyping revealed that the genetic diversity of Hib is relatively restricted in some populations (van Alphen *et al.*, 1983; Takala *et al.*, 1987; Clements *et al.*, 1992; Urwin *et al.*, 1995). It has also recently been reported that the same PCR ribotype was found among a majority of carrier Hib isolates from Aboriginal children living in a remote island community in the Northern Territory (Smith-Vaughan *et al.*, 1998).

Interestingly, the extent to which the predominant type, designated *Sma*I-M5, was present varied among both the metropolitan areas and the disease type subsets in this sample. Overall the predominant type alone accounted for one third of the meningitis isolates and one-third of the epiglottitis isolates. There was a difference in the number of epiglottitis isolates represented by the predominant type in the regional subsets. In the Canberra sample 7% of the isolates were the predominant type. Whereas in the Victorian sample of a similar size, 70% of the isolates were the predominant type. The proportion in Western Australia was almost half that of Victoria but still six times that of Canberra. So while it appears that the *Sma*I-M5 type was found equally among meningitis and epiglottitis isolates in the entire sample, anomalies were found among the urban subsets of epiglottitis isolates. One explanation is that the Victorian and Sydney samples were not representative proportions of the actual number of Hib types found in those areas.

There was less divergence in the representation of the predominant type in the subsets of meningitis isolates. It accounted for about 40% of the isolates in three of the urban subsets, the exception being the Canberra subset where it accounted for only 20%. The lower representation of the predominant type in the Canberra sample is not surprising as it was the most diverse of the urban subsets and had the highest ratio of types to isolates that was more than two times greater than that found in the Western Australian sample. The Victorian and Sydney ratios were similar and slightly higher than the Western Australian ratio of types to isolates.

No other type predominated in the combined sample and in the Canberra subset no type was represented by more than 2 isolates. Thus, no other single type had been selected for and, subsequently, accounted for the low percentage of the *Sma*I-M5 type in the Canberra sample.

It is notable that in the Victorian sample of epiglottitis isolates, 3 of the 4 types found were clonally related. The fourth type was less closely related to the others at an *F* value of 0.20 that was not outside the bounds of potentially clonally related isolates (Thal *et al.*, 1997). Thus, while 90% of the isolates were shown to belong to clonally related types, the entire sample was potentially clonally related. A similar homogeneity was not found in any other subset of epiglottitis isolates. However, a slightly greater homogeneity was found in the Western Australian subset of meningitis isolates where 94% of the isolates fell in the predominant type and its clonally related types. Even though the subset sample sizes are relatively small, the data suggest that there is a stronger selection in these subsets for the predominant type and a small group of its clonally related types. It is puzzling that this appears to occur among the epiglottitis isolates in Victoria and among the meningitis isolates in Western Australia and not in any other subsets of the sample.

The RFLPs of the 20 isolates from Victoria were compared with their previously determined OMP subtypes (Clements *et al.*, 1992). Eleven of the isolates were members of the same RFLP type (indistinguishable from *Sma*I-M5). Three OMP subtypes, 1L, 14L and 1VA were found among them. The most genetically distinct type, *Sma*I-V1, had two members with the same OMP subtype, 3L. Of the other 7 isolates that fell into 7 RFLP types, 6 were OMP subtype 1-VA and one was subtype 14L. As previously mentioned, when the fragment pattern of *Sma*I-M5 was compared to the *Sma*I RFLP

patterns of an OMP reference set of 21 isolates obtained from European and North American collections it was identical to the pattern of OMP subtype 1/3L called 1VA by Clements, *et al.* It might also be noted that OMP subtypes 1L and 14L had identical PFGE patterns that differed from the 1/3L (1VA) OMP subtype pattern by a two fragment difference. While the discriminatory ability of PFGE was significantly greater than that of OMP typing among the 20 Victorian isolates, in some cases at least, discrimination was better using both methods. The usefulness of correlating OMP subtyping with PFGE RFLP typing awaits further study.

A genetic distance showing a similarity of less than 20% was found between four types, M7-M13 and S11-S37 had F=0.19 and 0.20, respectively. Five types, M7, M9, M13, S37 and V1, representing 6 isolates, were found to be consistently less than 60% similar to all other types and mostly even more dissimilar to each other with the exception of M13 and S37 that were clonally related. In contrast the majority of the other types fell into closely related clusters. In fact, the predominant type and its clonally related types accounted for almost 70% of the sample. The presence of such diverse types in an otherwise relatively closely related group of types is intriguing. Do they represent endemic strains or strains recently brought into the region? Are they newly arisen strains or older strains from which the endemic clonal group has evolved?

No evidence was found to suggest that a particular genotype was associated with one disease and not the other. The data show clearly that the predominant type found in all urban areas and its clonally related types are associated with both meningitis and epiglottitis. Isolates associated with epiglottitis were found in 9 of 11 major clusters, including the most distant cluster, of the dendrogram constructed from the similarity matrix of the entire sample. Meningitis isolates were found in all the major clusters.

In summary, the population structure of Hib among meningitis and epiglottitis isolates from urban areas appears to be clonal. Typical of a clonal population, one type predominated and with its clonally related types accounted for almost 70% of the sample. A few distantly related types representing less then 6% of the sample were found. No association between RFLP type and disease state, age, sex, biotype, beta-lactamase production or geographic location was found. Hybridisation of the *Sma*I digested DNA fragments with pU082 did not split the types further than that found with PFGE alone.

CHAPTER 5

Investigation of the Genetic Diversity of 76 Aboriginal Isolates of *Haemophilus influenzae* type b Using Pulsed Field Gel Electrophoresis

5.1 Introduction

The pre-vaccine epidemiology of Hib disease among Aboriginal populations was markedly different from that found among non-Aboriginals in Australia. The incidence of disease was four to ten times higher, the age of onset of disease was earlier and while a spectrum of Hib diseases occurred, in particular, meningitis, pneumonia, and bacteremia, virtually no epiglottitis was found in Aboriginal populations (Hanna, 1992). Though generally kept low because the associated hospitalisation rates were very high, the case fatality rate was also higher (Hanna and Wild, 1991). This is similar to the epidemiology of Hib disease among other high-risk ethnic minorities, namely, American Indians (Losonsky *et al.*, 1984), Alaskan Eskimos (Ward *et al.*, 1981) and Canadian Inuits (Shapiro and Ward, 1991). The reasons for these variations are not understood. Environmental factors and genetic predisposition have been implicated but there are likely to be other factors involved, including differences in the distribution of Hib strains in different populations.

The annual incidence rate of Hib meningitis, the most precisely estimated disease incidence, was approximately 150 episodes per 100,000 Aboriginal children under five years of age per year in Western Australia and the Northern Territory (Hanna and Wild, 1991). However, the point estimate was much higher in Central Australia at 273 cases per 100,000 children (Hanna, 1990). In contrast the incidence of Hib meningitis among children in non-Aboriginal populations in the same time period was around 25/100,000.

As noted above, Central Australian Aboriginal children were found to have a significantly higher incidence of Hib disease then other Aboriginal children in the Northern Territory and Western Australia. In fact, the world's highest incidence of Hib disease was estimated to be a startling 991 cases per 100,000 Central Australian Aboriginal children and residence in Central Australia was identified as one of the major risk factors for invasive Hib disease (Hanna, 1990). Concurrently the incidence of meningitis in non-Aboriginal children in Central Australia was found to be four times greater than that estimated for Melbourne children (Hanna, 1990) reflecting the exceptionally high risk for Aboriginal children residing there. One of the factors that might explain the increased risk of invasive Hib infection in Central Australia is the possibility of the presence of a more virulent strain than that prevalent in other areas.

It can also be stated that Hib caused a substantial proportion of pneumonia and bacteraemia in young Aboriginal children and less common but severe infections such as septic arthritis, cellulitis and osteomyelitis. Not a single case of epiglottitis among Aboriginal children was documented in the years covered by the three major studies on the epidemiology of Hib disease among Aboriginal populations in Western Australian and the Northern Territory (Hanna, 1990; Hanna and Wild, 1991; Hanna *et al.*, 1992).

There is little information about the diversity of Hib isolates among Aboriginals, but there is some intriguing evidence pointing to a difference among Aboriginal and non-Aboriginal isolates. When outer membrane protein (OMP) subtyping was used to distinguish 220 non-Aboriginal isolates from Melbourne and Sydney and 55 rural Aboriginal isolates from the Northern Territory and Western Australia, a marked discrepancy between OMP subtypes causing Hib disease in the Aboriginal and non-Aboriginal children was found (Gilbert and Clements, 1993). It appeared that the predominant subtype (1NT) among Aboriginal children was similar, but distinguishable from, the predominant Victorian/Sydney subtype (1VA) and that there was more variation in subtypes, some of which appeared novel among the Aboriginal isolates.

To begin to unravel the mystery concerning the epidemiology of Hib disease among Aboriginal populations it would be particularly interesting to determine if:

1) the difference between Aboriginal and non-Aboriginal OMP subtypes detected by Gilbert *et al.* (1993) is reflected in PFGE RFLP types,
2) the *Sma*I-RFLP type predominating in the non-Aboriginal meningitis and epiglottitis sample described in Chapter 4 is found among Aboriginal isolates or if other unique and potentially more virulent types predominate, and
3) the Aboriginal isolates in Central Australia are different from those found in the Top End of the Northern Territory and Western Australia.

The purpose of this study was to use PFGE to examine the molecular genetic characteristics of Hib isolates from Aboriginal patients living in distinct geographic regions to determine if there are genetically detectable differences between them and, subsequently, to determine of there are differences between Aboriginal and non-Aboriginal isolates.

5.2 Methods

Preparation of DNA, restriction endonuclease digestion, PFGE conditions, determination of the numbers and mobilities of fragments, estimation of genetic diversity and construction of dendrograms and hybridisation were performed as previously described.

5.2.1 Bacterial isolates (Table 2.1, Appendices A.4-A.6)

The sample comprised 76 isolates from Aboriginal patients and carriers living in the Alice Springs region, Bathurst Island, rural Western Australia and metropolitan Perth. It included 23 isolates recovered from patients with meningitis, 29 isolates from other diagnoses and 24 isolates recovered from carriers. No isolates associated with epiglottitis in Aboriginals were found. All the isolates were collected between 1985 and 1993. Thirty-six were obtained from Westmead Hospital, Westmead, New South Wales from a collection originally obtained from Queensland Institute of Medical Research, Brisbane, Queensland; 9 were obtained from the Royal Children's Hospital, Melbourne, Victoria; 19 were obtained from the Menzies School of Health Research, Darwin, Northern Territory and 12 from were obtained from the Princess Margaret Hospital, Perth, Western Australia. A breakdown of the geographic location and disease association of the isolates is shown in Table 2.1. Appendices A.4-A.6 describe the isolates in more detail.

All of the isolates obtained for this study had been identified by conventional methods as described previously. They were stored at –70°C or, in the case of the isolates from Western Australia, were lyophilised. *Sma*I was used for the analysis of RFLPs in this study, however, digestion with *Apa*I was also performed and the results were available if needed.

The isolates were designated as follows:

Source	Designation	Reason for breaks in numbering of isolates
Alice Springs region (*n=45*) [a]	A1-A50 excluding A10, A24, A25,A30, A48	A10 and A24 were from non-Aboriginal patients A25 was a nontypeable *Haemophilus influenzae* A30 and A48 were each members of a pair of isolates from one patient with indistinguishable RFLPs.
Bathurst Island (*n=19*) [b]	B1-B21 excluding B9 and B17	B9 and B17 were non-viable on receipt.
Rural Western Australia (*n=7*) [c]	WA2, 3, 5, 6, 7, 9, 10	WA1, 4, 8, 12, and 13 are from metropolitan Perth; the designation WA11 was excluded because the isolate was non-viable on receipt.
Perth (*n=5*) [c]	WA1, 4, 8, 12, 13	WA2, 3, 5, 6, 7, 9, and 10 are from rural Western Australia.

[a]From collections held at Westmead Hospital, Westmead, New South Wales originally obtained from the Queensland Institute of Medical Research, Brisbane, Queensland (*n=36*) and Royal Children's Hospital, Melbourne, Victoria (*n=9*)
[b]From a collection held at Menzies School of Health Research, Darwin, Northern Territory
[c]From a collection held at Princess Margaret Hospital, Perth, Western Australia

5.3 Results

Initially, 46 isolates were obtained from the Alice Springs region but one failed to hybridise with either pU038 or pU082 and was excluded from the study. All the other isolates in this study hybridised with pU082 but no further splitting of types was detected by *cap* polymorphisms. All the isolates from the Alice Springs region (*n=45*) hybridised with the P2 probe but no further splitting of types was detected. The visual analysis of the RFLP patterns is described first and then the numerical analysis is described.

5.3.1 Visual analysis of *Sma*I-RFLPs

Tables 5.1-5.3 describe the fragment distribution among types found in each subset. Fragment sizes ranged from ~6 kb to 550 kb. While this extended the size range from that found among 104 non-Aboriginal isolates, the size of fragments was well within the ability of the system to detect them using the ramping configuration previously determined with non-Aboriginal isolates. There was no need to adjust methods to evaluate the Aboriginal isolates.

Alice Springs region, Northern Territory Aboriginal isolates (Figures 5.1-5.5, Table 5.1) The 45 Aboriginal isolates from the Central Australia Alice Springs region were associated with meningitis (*n=11*) pneumonia (*n=9*), gastroenteritis (*n=9*), acute lower respiratory tract infection (*n=4*), cellulitis (*n=1*), febrile disease (*n=1*), failure to thrive (*n=1*), conjunctivitis (*n=1*), bronchiolitis (*n=1*) and otitis media (*n=1*). Six isolates were recovered from children that did not exhibit any sign of disease. It is not known whether they subsequently developed Hib disease.

Thirteen genetic types were found. One of them, A1, accounted for 49% of the sample. A single isolate, A22, was indistinguishable from the predominant type (*Sma*I-M5) previously found among non-Aboriginal urban isolates. The distribution of isolates among types was:

Northern Territory Alice Springs region isolates (*n=45*)

*Sma*I-RFLP type	Number of isolates/type
A1	22
A5	6
A23	5
A2	3
A3, A8, A21, A22*, A34, A41, A42, A45, A47	1 each

*indistinguishable from RFLP type *Sma*I-M5

The distribution of types among meningitis isolates, other diagnoses, and carriers was not restricted. The predominant type, A1, was found among the major disease states, meningitis (*n=4*), pneumonia (*n=4*) and gastroenteritis (*n=4*), and the less common serious diseases (n=6) and among the carrier isolates (*n=4*). The second most frequent type, A5, was found among the carrier isolates (*n=1*) and meningitis (*n=1*), pneumonia (*n=2*), and acute lower respiratory tract infections (*n=1*) but none was found among the 9 gastroenteritis isolates.

Pairs of isolates for each of four patients were included in this sample. Two pairs had members that shared the same RFLP type. The two other pairs each had members with different RFLPs; but, in each case the RFLPs of a pair with different types were closely related differing by only 2 or 4 fragments. The distribution of the types among the 4 pairs was:

RFLP typing of paired isolates from a single patient

Diagnosis	Pair	Source	Type	Fragments shared
Meningitis	A21 A22	blood csf	A21 A22	24 of 28
Meningitis	A34 A35	blood csf	A34 A1	30 of 32
Pneumonia	A29 A30	blood pleural fluid	A1 A1	32 of 32
Meningitis/ otitis media	A47 A48	not known not known	A47 A47	32 of 32

Bathurst Island, Northern Territory Aboriginal isolates (Figures 5.6-5.7, Table 5.2) This sample consisted of 1 isolate recovered from an ear discharge in an infant and 18 recovered from the nasopharynx of infants who were not exhibiting symptoms of Hib disease. Six RFLP types were found. No single type predominated, but two types had 7 and 6 members each. One of these, *Sma*I-B6, was indistinguishable from *Sma*I-A1, the predominant type in the Alice Springs region. The predominant type (*Sma*I-M5) found among the urban isolates was not represented in this sample. The distribution of the isolates among the types was:

Northern Territory Bathurst Island isolates (*n=19*)

*Sma*I-RFLP type	Number of isolates/type
B2	7
B6	6
B8	3
B1, B5, B13	1 each

The type designated B13 represented the isolate recovered from an ear discharge. It was clonally related to type B2 (*n=7*) sharing 28 of 30 fragments found between them. Three sets of 2 isolates recovered at different times from three children and one set of 3 isolates collected on the same date from one child were included in the sample. The RFLPs found were indistinguishable in each respective set. The date of collection and distribution of types among these sets was:

Date of Collection	Set	Source	Type
14 April 92	B2	nasopharynx	B2
14 April 92	B3	nasopharynx	B2
14 April 92	B4	nasopharynx	B2
21 July 92	B6	nasopharynx	B6
02 Dec 92	B7	nasopharynx	B6

02 Dec 92	B14	nasopharynx	B6
10 Feb 93	B15	nasopharynx	B6
10 Aug 93	B18	nasopharynx	B2
20 Sept 93	B19	nasopharynx	B2

Western Australian Aboriginal isolates (Figures 4.6, 5.8, Table 5.3) Twelve meningitis isolates from Aboriginal children living in metropolitan Perth (*n=5*) and rural Western Australia (*n=7*) were examined. The rural isolates were from a relatively large geographic area including the Kimberly, Midlands, Central and Southeast Western Australia.

Five RFLP types were found. Seventy-five percent of the sample was represented by two clonally related types, WA4 and WA1, that shared 30 of 32 fragments between them. WA4 was indistinguishable from the predominant type, A1, in the Alice Springs region. It comprised 7 (64%) of the sample including 3 urban isolates and 4 rural isolates while WA2 comprised 2 rural isolates. One type, WA1, representing a single isolate was indistinguishable from *Sma*I-M5. The distribution of isolates to types was:

Metropolitan Perth isolates (*n=5)*		Rural Western Australia isolates (*n=7)*	
*Sma*I-RFLP type	Number of isolates/type	*Sma*I-RFLP type	Number of isolates/type
WA4	3	WA4	4
WA1*, WA12	1 each	WA2	2
		WA5	1

*indistinguishable from RFLP type *Sma*I-M5

All Aboriginal isolates (Table 5.4) A total of 19 types were found among the 76 Aboriginal isolates. One type, designated A1 (*n=36*), predominated and was the only one found in all three subsets. It was clonally related to A34 (*n=1*), B5 (*n=1*) and A23 (*n=7*). These related types accounted for 59% of the sample. Five other types had multiple members. Type A5 was the most diverse type and most distantly related to A1 with whom it shared only 2 of 31 fragments. A summary of the distribution of types is shown in Table 5.5.

Table 5. 1 Number of *Sma*I-RFLP types and the number of fragments found per type among 45 Aboriginal isolates of Hib from the Alice Springs region

	Total number of types		Number of fragments per type 13	14	15	16	17
Meningitis isolates *n=11*	7 (0.64:1)[a]	Types	2	1	2	2	
		Isolates	*n=2*	*n=1*	*n=3*	*n=5*	
Other isolates *n=28*	9 (0.32:1)[a]	Types	0	1	4	3	1
		Isolates	*n=0*	*n=1*	*n=6*	*n=20*	*n=1*
Carrier isolates *n=6*	3 (0.50:1)[a]	Types	0	1	1	1	
		Isolates	*n=0*	*n=1*	*n=1*	*n=4*	
All isolates *n=45*	13 (0.29:1)[a]	Types	2	1	5	4	1
		Isolates	*n=2*	*n=3*	*n=10*	*n=29*	*n=1*

[a]Ratio of the number of types found to the total number of isolates typed.

Table 5. 2 Number of *Sma*I-RFLP types and the number of fragments found per type among 19 Aboriginal carrier isolates of Hib from Bathurst Island

	Total number of types		Number of fragments per type 14	15	16
All isolates *n=19*	6 (0.32:1)[a]	Types	1	3	2
		Isolate	*n=1*	*n=11*	*n=7*

[a]Ratio of the number of types found to the total number of isolates typed.

Table 5. 3 Number of *Sma*I-RFLP types and the number of fragments found per type among 12 Aboriginal isolates of Hib from Western Australia

	Total number of types		Number of fragments per type 13	14	15	16
Rural isolates *n=7*	3 (0.64:1)[a]	Types	0	1	0	2
		Isolates	*n=0*	*n=1*	*n=0*	*n=6*
Perth isolates *n=5*	3 (0.60:1)[a]	Types	1	1	0	1
		Isolates	*n=1*	*n=1*	*n=0*	*n=3*
All isolates *n=12*	5 (0.42:1)[a]	Types	1	2	0	2
		Isolates	*n=1*	*n=2*	*n=0*	*n=9*

[a]Ratio of the number of types found to the total number of isolates typed.

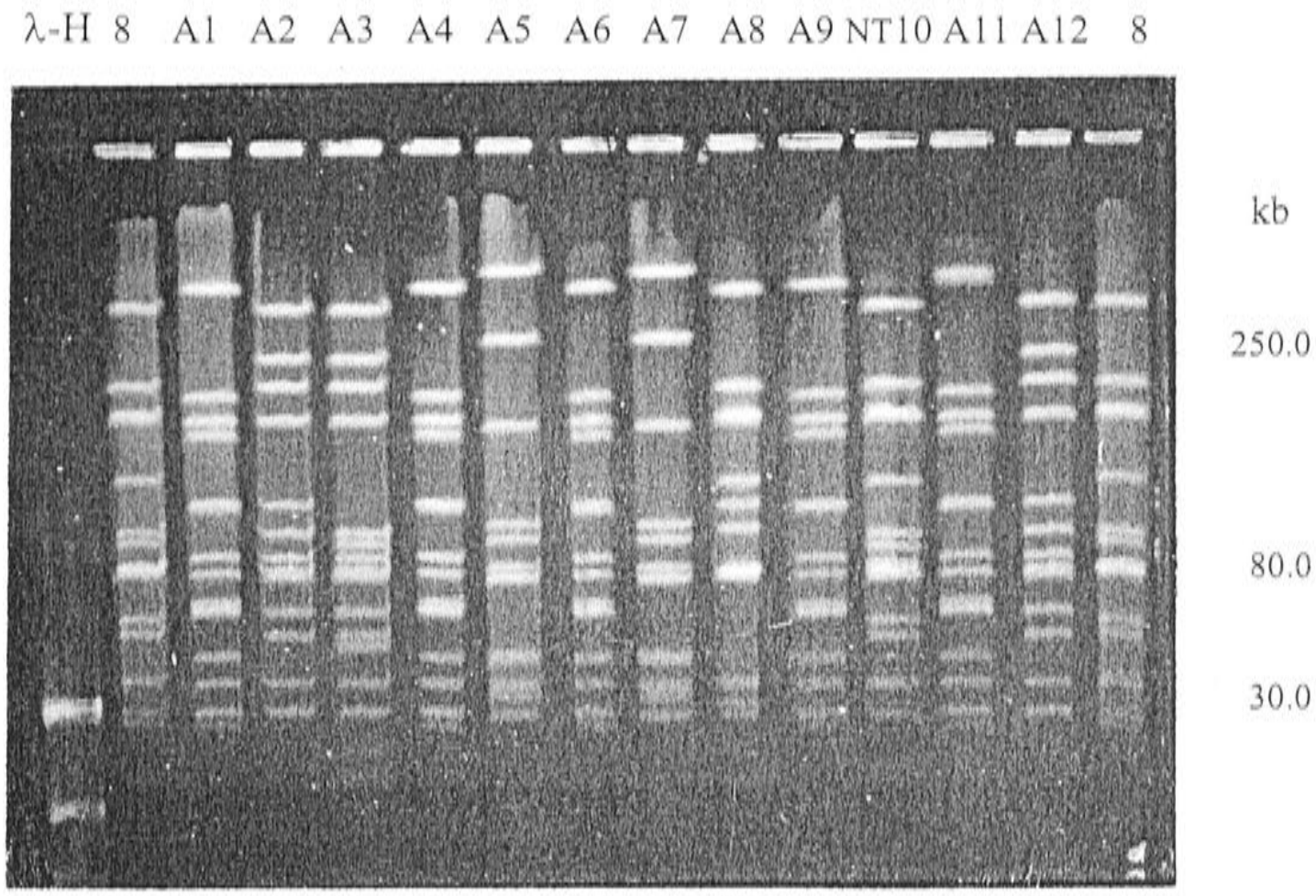


Figure 5.1 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 11 Aboriginal Hib isolates and 1 non-Aboriginal isolate from the Alice Springs region. Lanes labelled 8 and λ-H contain fragment standards.

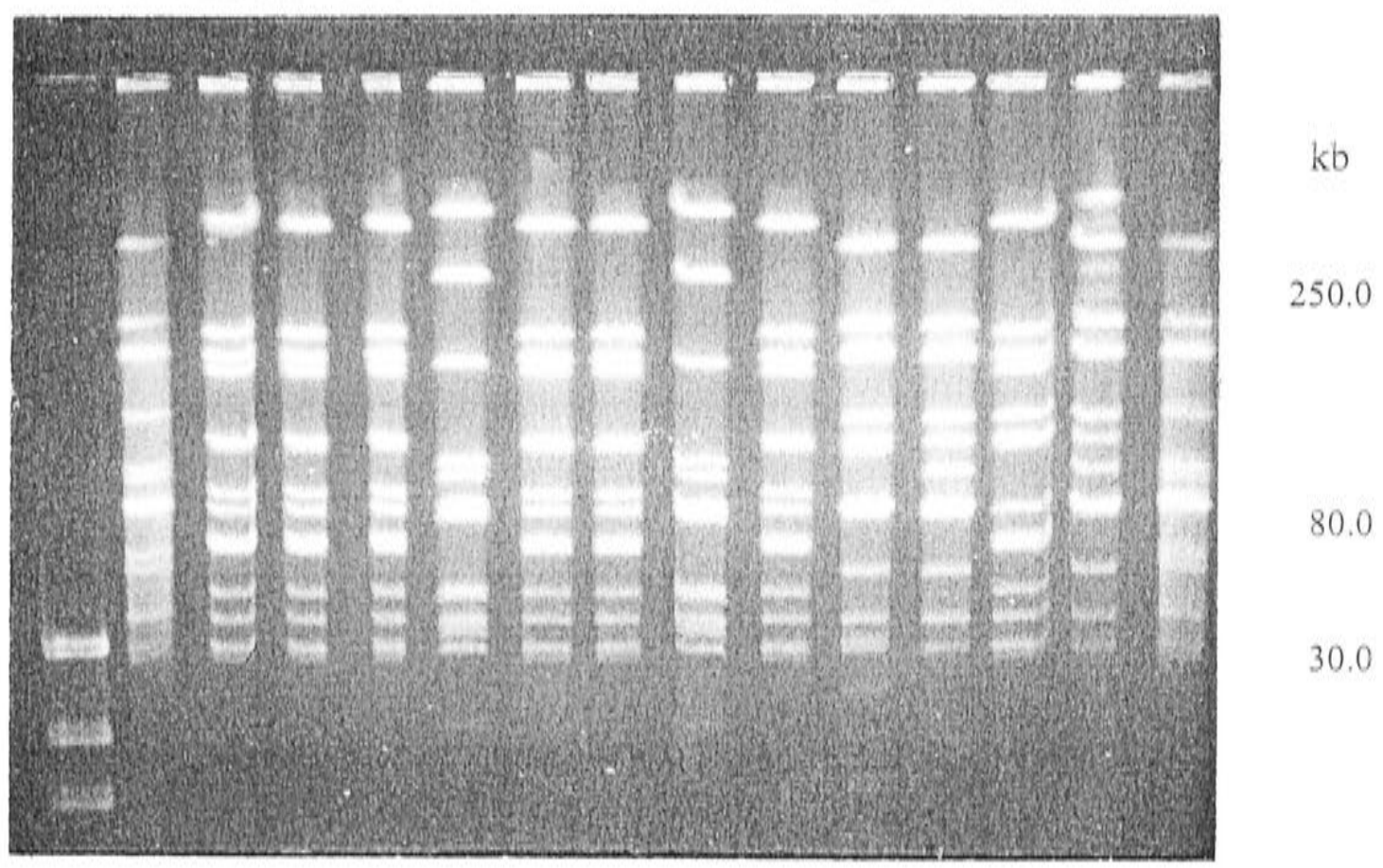


Figure 5.2 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 12 Aboriginal Hib isolates from the Alice Springs region. Lanes labelled 8 and λ-H contain fragment standards.

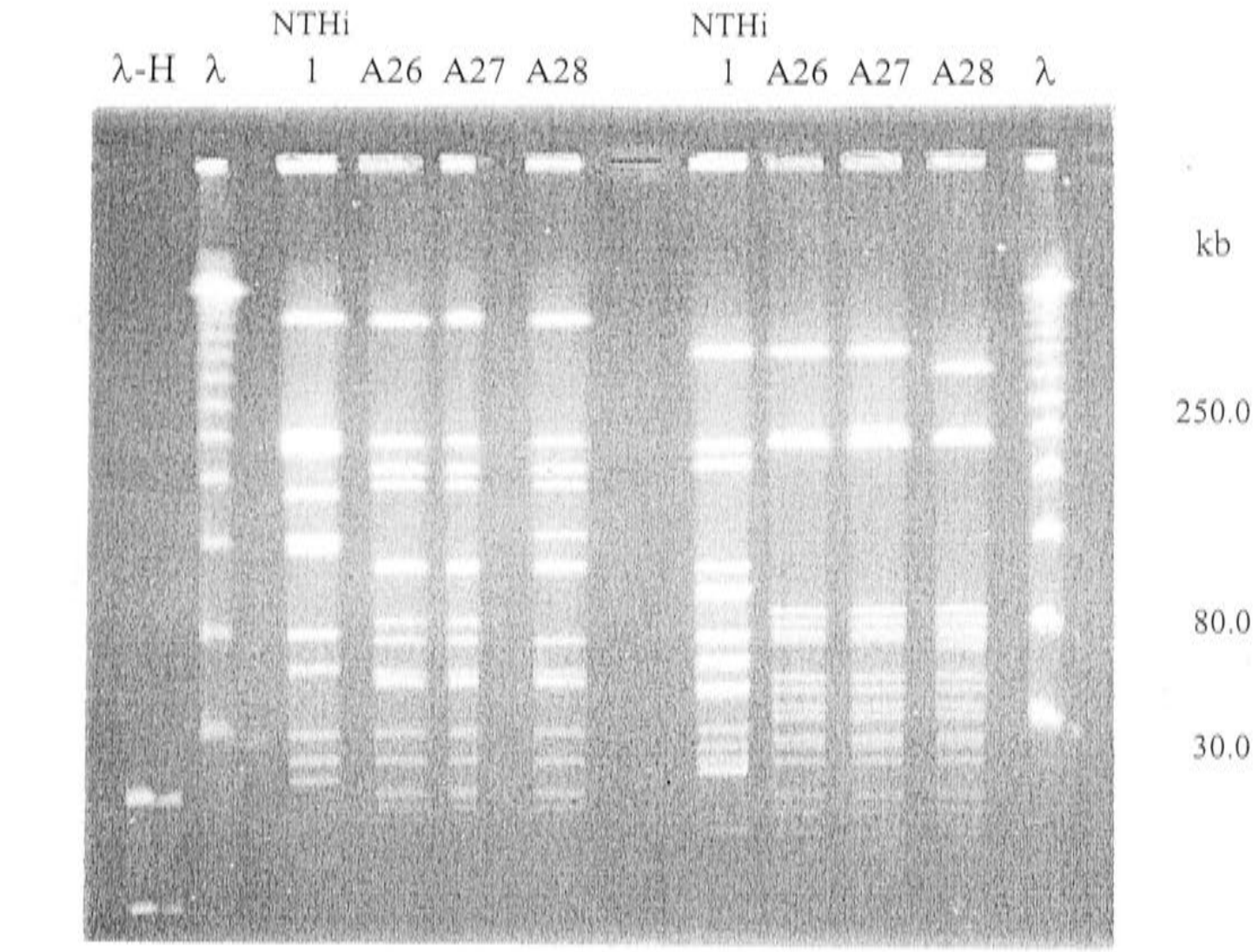


Figure 5.3 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 3 Aboriginal Hib isolates and 1 nontypeable *H. influenzae* isolate from the Alice Springs region. Lanes labelled λ and λ-H contain fragment standards.

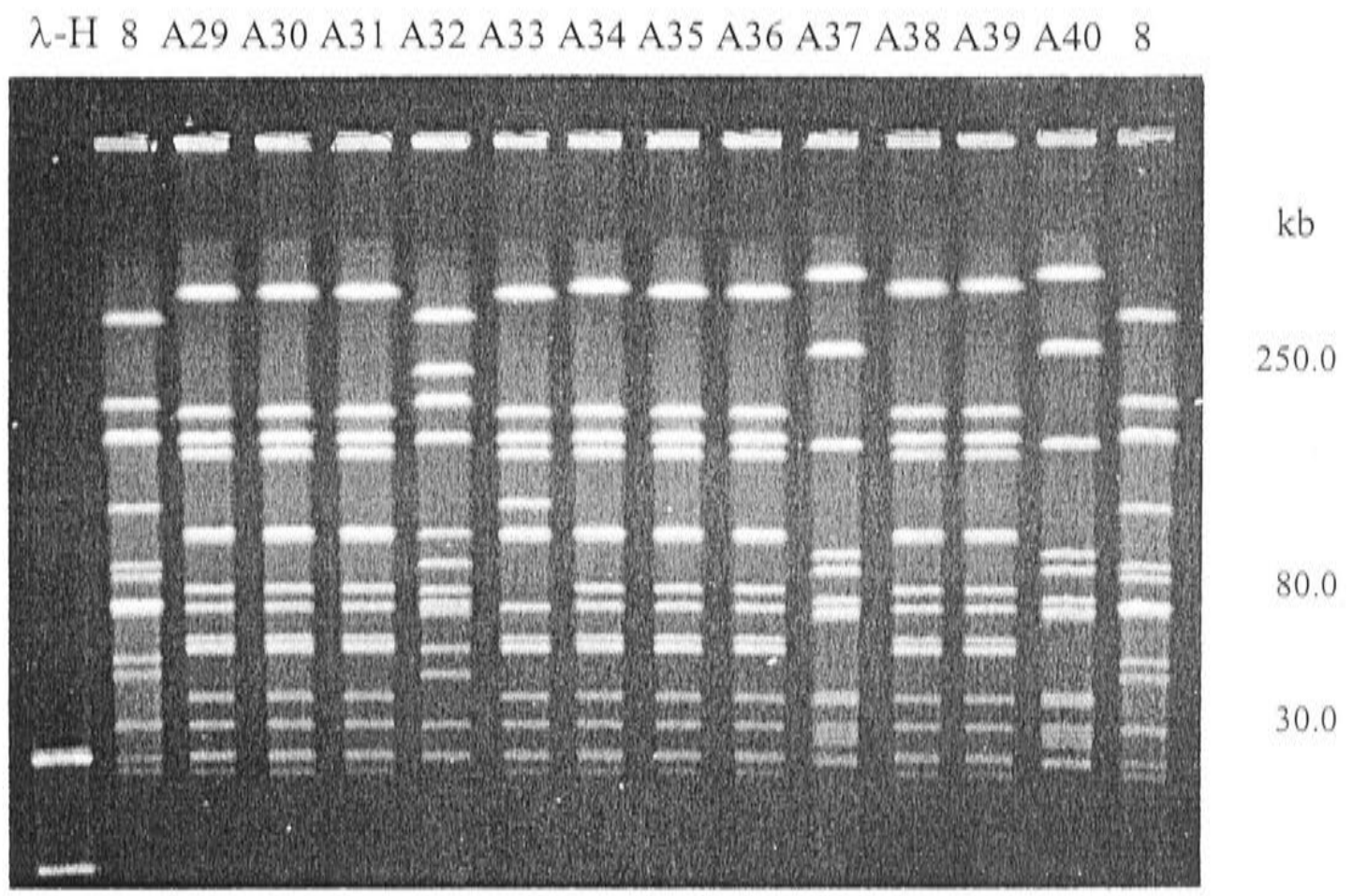


Figure 5.4 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 12 Aboriginal Hib isolates from the Alice Springs region. Lanes labelled 8 and λ-H contain fragment standards.

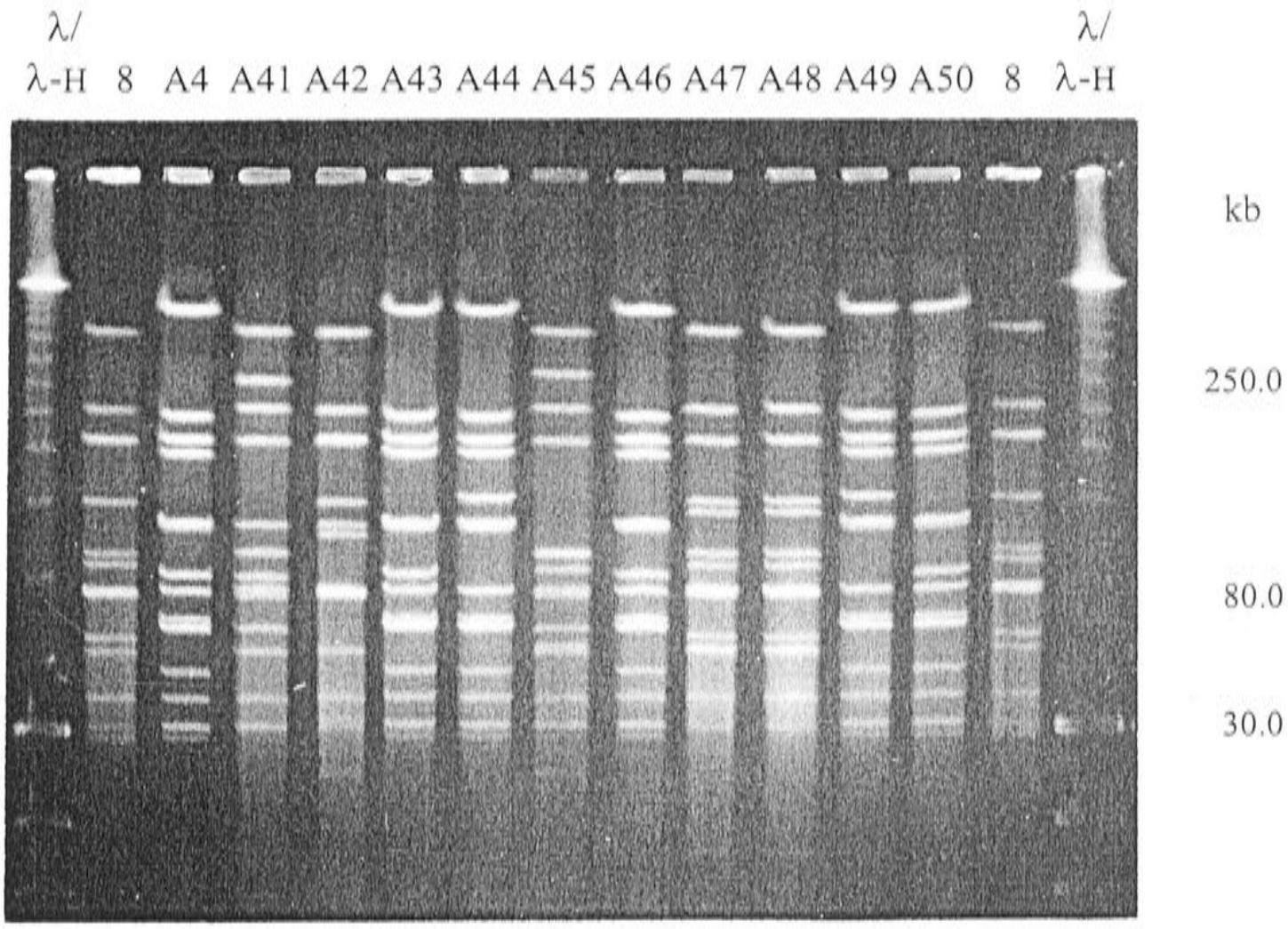


Figure 5.5 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 11 Aboriginal Hib isolates from the Alice Springs region. Lanes labelled 8 and λ/λ-H contain fragment standards.

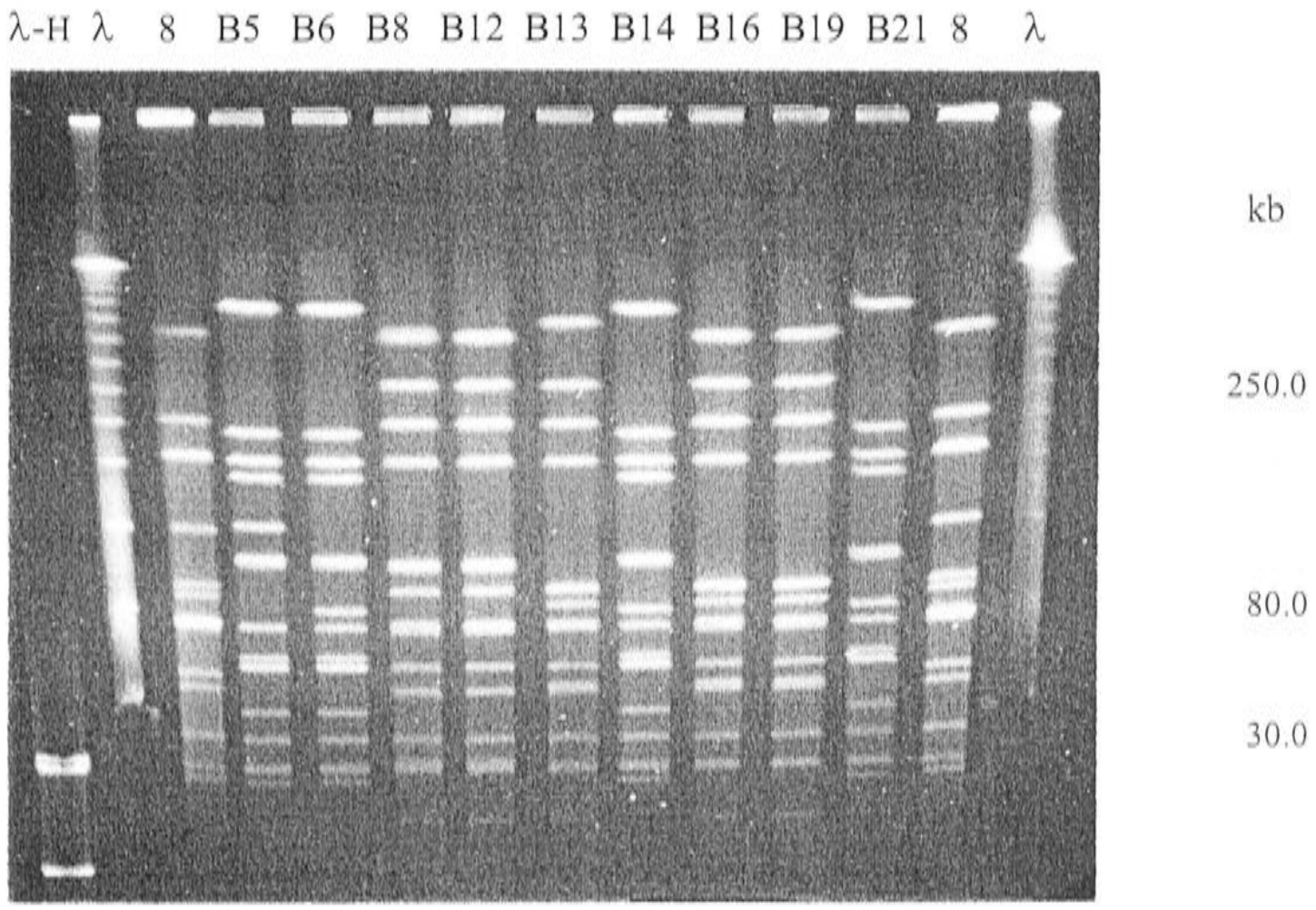


Figure 5.6 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 9 Aboriginal Hib isolates from Bathurst Island. Lanes labelled 8, λ and λ-H contain fragment standards.

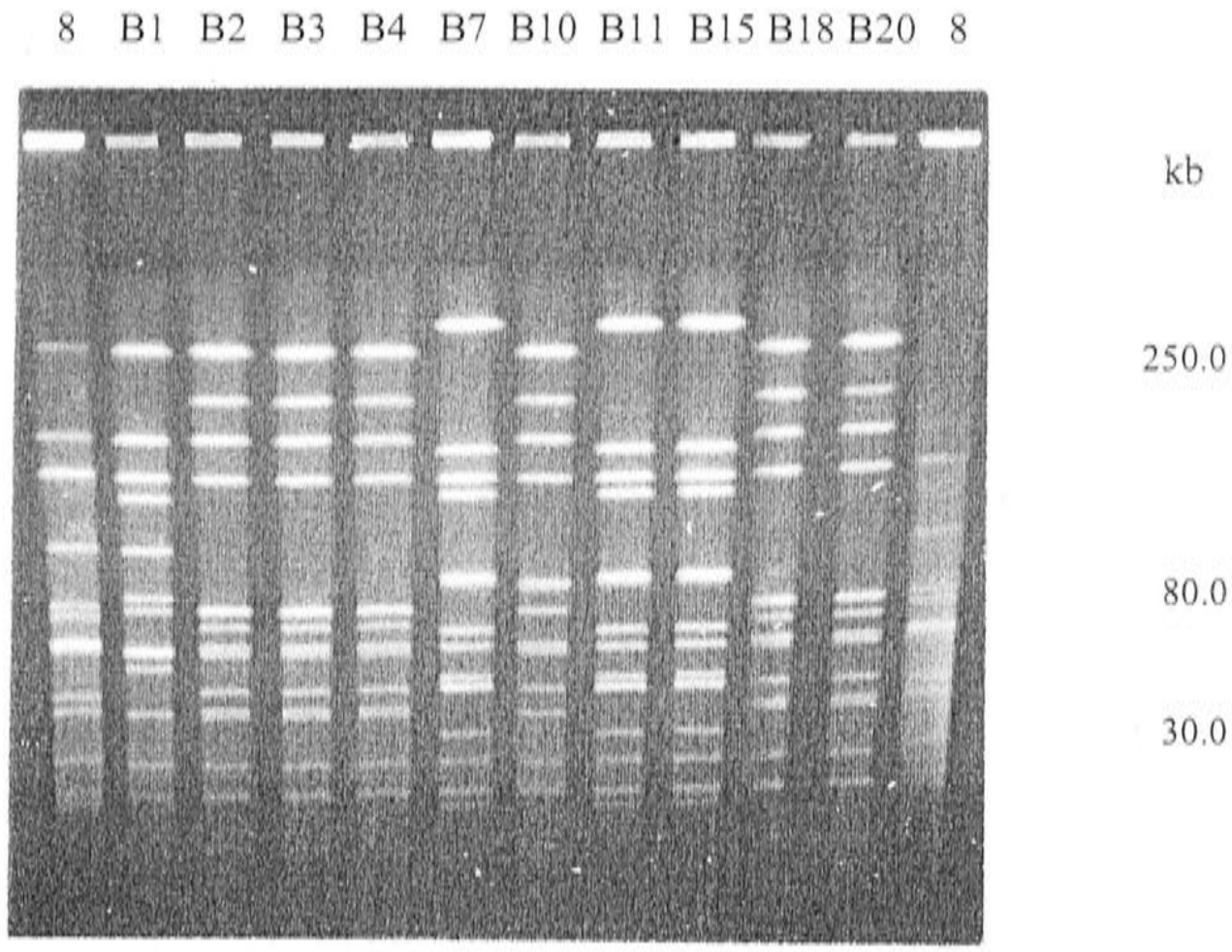


Figure 5.7 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 10 Aboriginal Hib isolates from Bathurst Island. The lane labelled 8 contains fragment standards.

λ/
λ-H 8 WA5 WA6 WA7 WA8 WA9 WA10 A1 8 λ

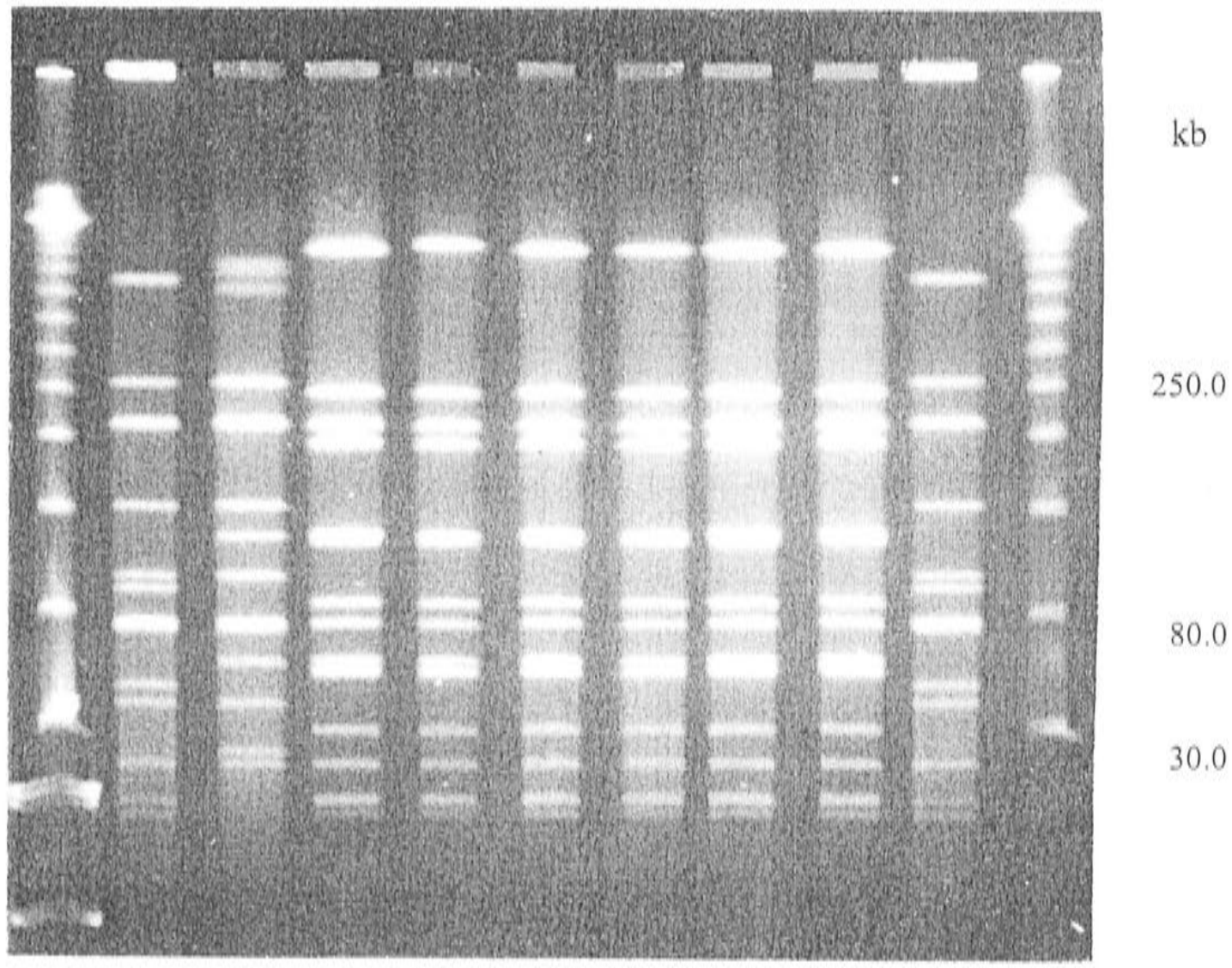


Figure 5.8 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 6 Aboriginal Hib isolates from Western Australia and 1 from the Alice Springs region. Lanes labelled 8, λ and λ-H contain fragment standards.

5.3.2 Genetic relationships determined by numerical analysis of restriction fragment length polymorphisms

The range of similarity or genetic distance, identification of clonal and the most diverse relationships among types and construction of dendrograms were determined from matrices produced using the mathematical model of Nei and Li as previously described.

Alice Springs region, Northern Territory Aboriginal isolates (Figure 5.9) Similarity among the types found in this sample ranged from 13-96%. Only 4 clonal pairs were among them. The predominant type, A1 (*n=22*), was clonally related to A34 (*n=1*) and A23 (*n=5*). Together this cluster of related isolates accounted for 62% of the sample (*n=45*). Interestingly no types in this cluster were closely related to other types. Their similarity ranged from 13-39% when compared to the other types in the sample. RFLP type A5 was the most genetically diverse. It was less then 21% similar to all other types and as low as 13% similar when compared to A1, A23, A34 and A47.

Bathurst Island, Northern Territory Aboriginal isolates (Figure 5.10) The range of similarity among these types was 27 to 94% with two clonal pairs among them. No one type predominated but two distinct types that differed at $F \leq 0.50$ accounted for 79% of the isolates. One of these, B6 (*n=6*), was clonally related to B5 (*n=1*) and together they comprised 37% of the sample (*n=19*). The other, B2 (*n=7*) was clonally related to B13 (*n=1*) and they accounted for 42% of the sample.

Western Australian Aboriginal isolates (Figure 5.11) The range of similarity among the five types in this sample was 33-94%. RFLP type WA4 (*n=7*) was clonally related to WA2 (*n=2*). They comprised 75% of the sample and were the only clonally related pair of types found. They were genetically distinct from the other three types differing from them at an F value ≤ 0.41. One of these types, WA1 (*n=1*), was indistinguishable from *Sma*I-M5.

All Aboriginal isolates (Figure 5.12, Table 5.4) The 76 Aboriginal isolates fell into 4 major clusters two of which were genetically distinct from each other and the other two clusters at $F \leq 0.45$. The range of similarity was 13-96% with a 0.25:1 ratio of types to isolates. The predominant type, A1 (*n=36*) was clonally related to A34 (*n=1*), A23 (*n=7*) and B5 (*n=1*) and together they comprised 59% of the sample. The most

genetically distinct type, A5 (*n=6*), was found only in the Alice Springs region subset but otherwise there was no obvious restriction of types within the major clusters though only type A1 was found in all three subsets. Only 6 clonal pairs were found among the 19 types and four pairs (A5-A1, A5-A23, A5-A34, A5-47) were the most genetically diverse with an $F = 0.13$. A summary of the distribution of types is shown in Table 5.4.

Table 5. 4 Summary of *Sma*I-RFLP types* found among 76 Aboriginal isolates

Alice Springs *n=45*		Bathurst Island *n=19*	Western Australia *n=12*	All isolates *n=76*	
A1[a]	**A23**[d]	B1	WA1	**A1**[a,f,i]	A42
A2[b]	A34	**B2**[e]	**WA2**[h]	A2	**A45**[e,j]
A3	A41	B5	**WA4**[i]	A3	A47
A5[c]	A42	**B6**[f]	WA5	**A5**[c]	B1
A8	A45	**B8**[g]	WA12	A8	B5
A21	A47	B13	*n=5*	A21	**B8**[g]
A22	*n=13*	*n=6*		A22	B13
				A23[d,h]	WA5
				A34	WA12
				A41	*n=19*

[a]A1,A4,A6,A9,A11,A13, A14,A15,A17,A18,A20, A26,A27,A29,A31,A35, A36,A38,A39,A43,A46, A50 (*n=22*)
[b]A2,A12,A32
[c]A5,A16,A19,A37,A40
[d]A23,A28,A33,A44,A49
[e]B2,B3,B4,B16,B18,B19, B20
[f]B6,B7,B11,B14,B15,B21
[g]B8,B10,B12
[h]WA2,WA3
[i]WA4,WA6,WA7,WA8, WA8,WA10,WA13
[j]A45

*The designated *Sma*I-RFLP type name represents the lowest number assigned to an isolate of that type found among the Alice Springs isolates first, then new types found among the Bathurst Island isolates, and finally new types found among the Western Australian isolates. The types in boldface represent types with multiple members. The boxed type represents the one type found among all the subsets.

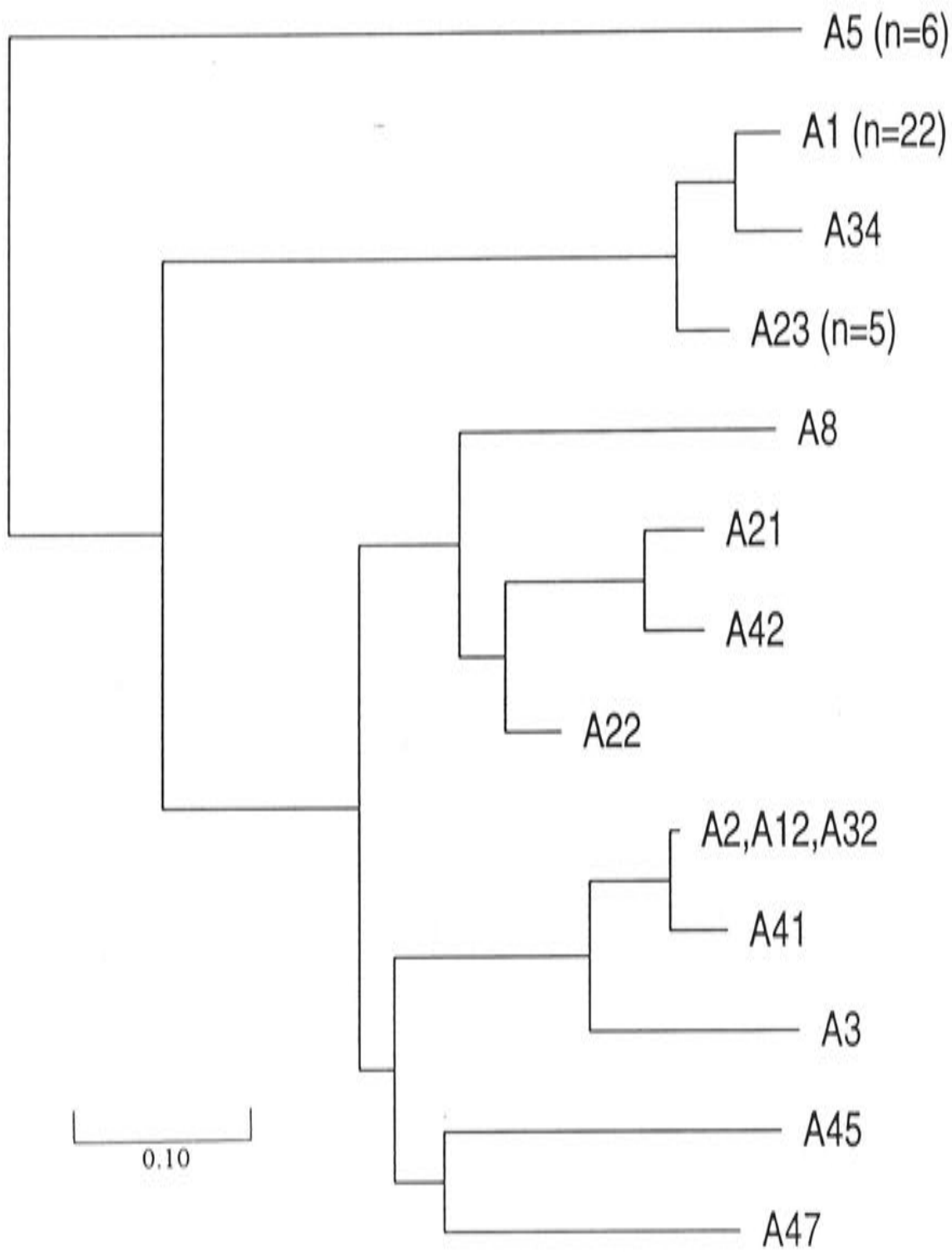


Figure 5.9 **Dendrogram showing the clustering of 13 *Sma*I-RFLP types found among 45 isolates obtained from Aboriginal patients and carriers in the Alice Springs region, Northern Territory.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbour-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the horizontal lines between pairs is an estimate of the genetic distance between the two types.

Types that fell in clusters ($n=3$) at $F \leq 0.5$	Types that fell in the clusters or branches ($n=10$) at $F \geq 0.9$	Pairs of types ($n=4$) at $F \geq 0.9$
1) A5	1) A1[a], A34 6) A2, A41	A1-A23
2) A1, A34, A23	2) A23[b] 7) A3	A1-A34
3) All the other types	3) A8 8) A45	A2-A41
	4) A21, A42 9) A47	A21-A42
	5) A22[c] 10) A5[d]	

[a]This *Sma*I-RFLP type included isolates A1, A4, A6, A9, A11, A13, A14, A15, A17, A18, A20, A26, A27, A29, A31, A35, A36, A38, A39, A43, A46 and A50 ($n=22$).

[b]This *Sma*I-RFLP type included isolates A23, A28, A33, A44 and A49 ($n=5$).

[c]This type was indistinguishable from *Sma*I-M5, the predominant type found among 104 non-Aboriginal isolates.

[d]This *Sma*I-RFLP type included isolates A5, A7, A16, A19, A37 and A40 ($n=6$).

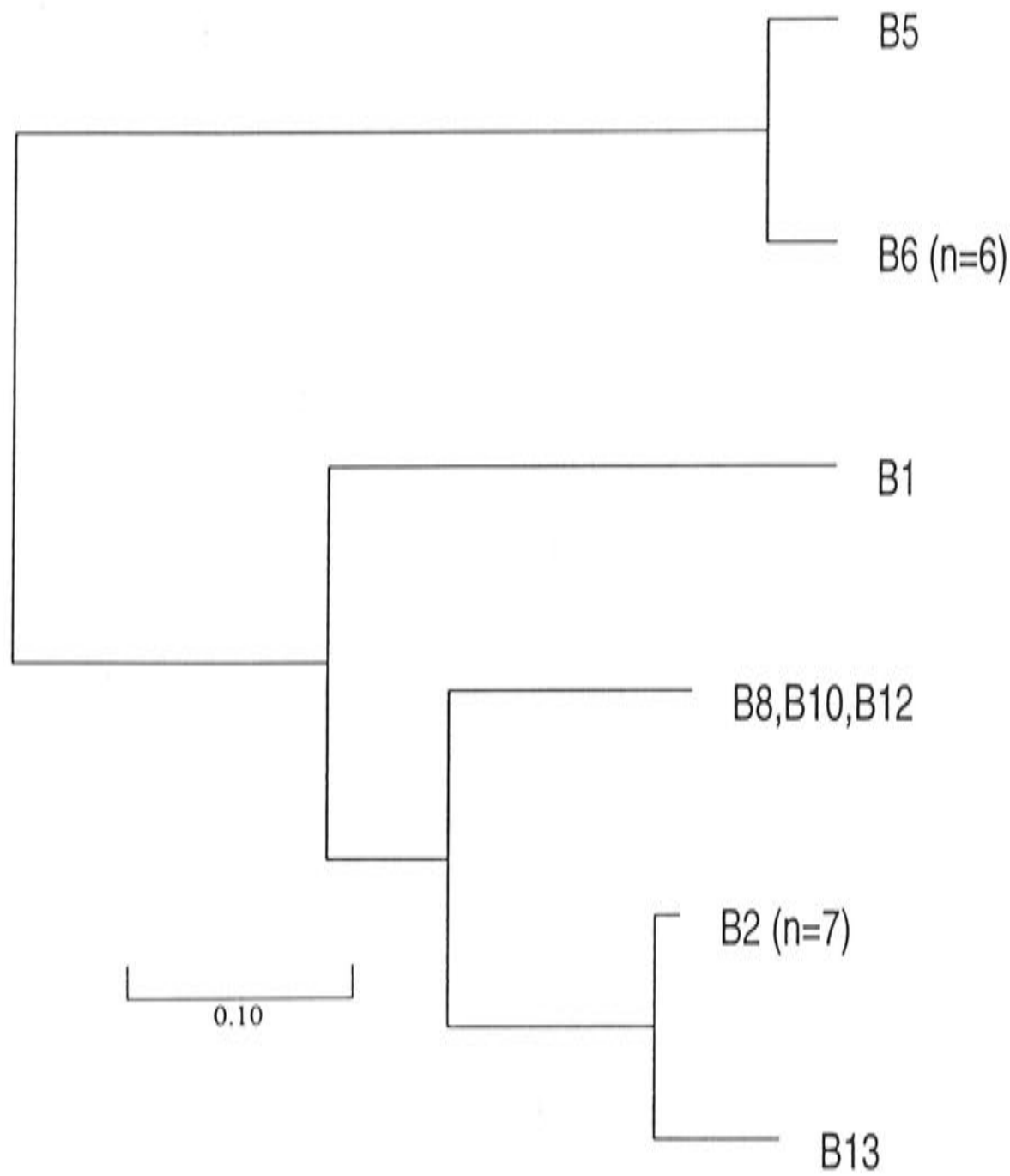


Figure 5.10 **Dendrogram showing the clustering of 6 *Sma*I-RFLP types found among 19 isolates obtained from Aboriginal children from Bathurst Island, Northern Territory.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbour-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the horizontal lines between pairs is an estimate of the genetic distance between the two types.

Types that fell in clusters ($n=2$) at $F \leq 0.5$	Types that fell in the clusters or branches ($n=4$) at $F \geq 0.9$	Pairs of types ($n=2$) at $F \geq 0.9$
1) B5, B6	1) B5, B6[a]	B2-B13
2) All the other types	2) B1	B5-B6
	3) B8	
	4) B2[b]	

[a]This *Sma*I-RFLP type included isolates B6, B7, B11, B14, B15 and B21 ($n=6$).
[b]This *Sma*I-RFLP type included isolates B2, B3, B4, B16, B18, B19 and B20 ($n=7$).

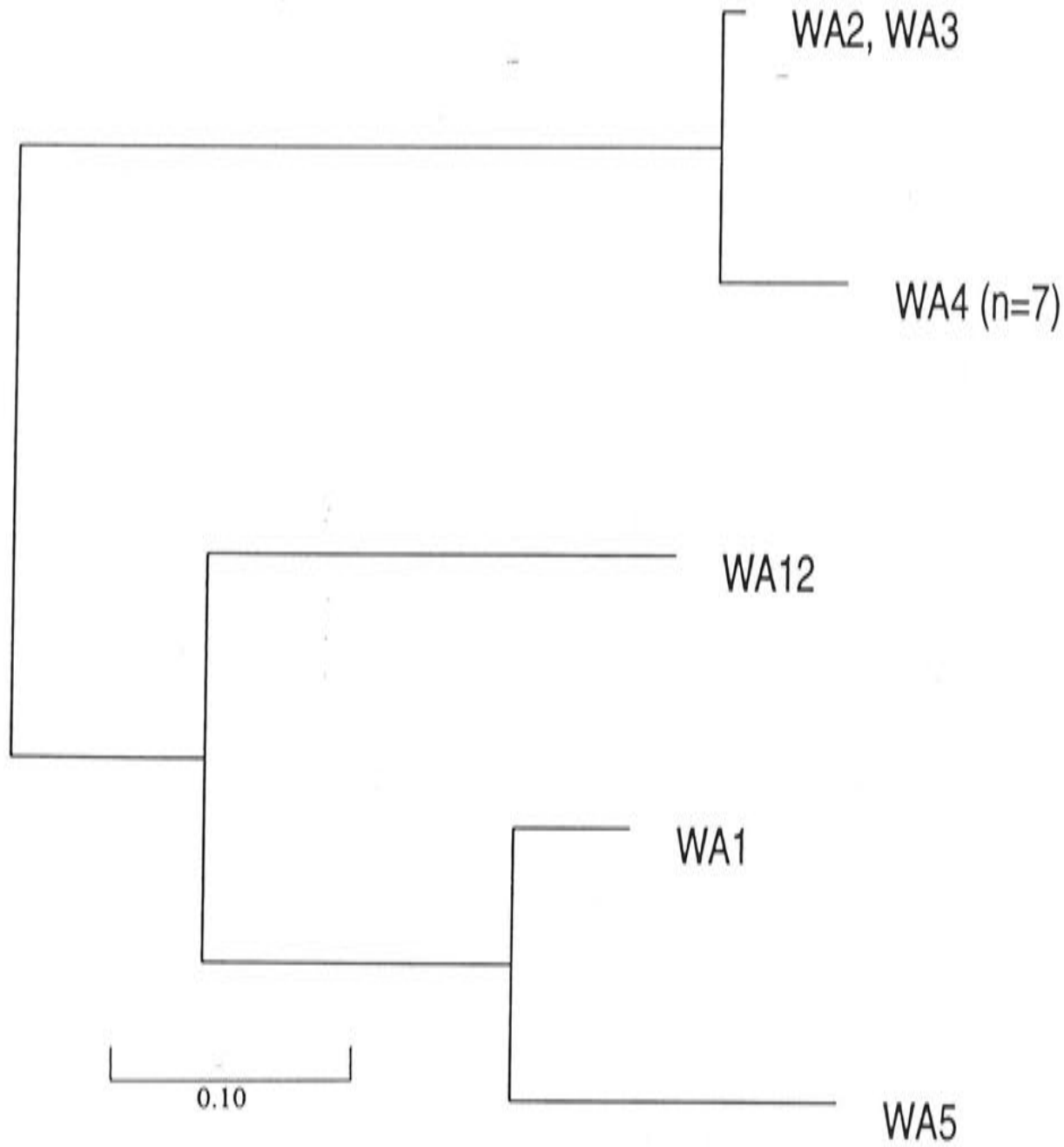


Figure 5.11 **Dendrogram showing the clustering of 5 *Sma*I-RFLP types found among 12 isolates obtained from Aboriginal patients in Western Australia.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbour-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the horizontal lines between pairs is an estimate of the genetic distance between the two types.

Types that fell in clusters ($n=2$) at $F \leq 0.5$	Types that fell in the clusters or branches ($n=4$) at $F \geq 0.9$	Pairs of types ($n=1$) at $F \geq 0.9$
1) WA2, WA4	1) WA2, WA4[a]	WA4-WA2
2) All the other types	2) WA12	
	3) WA1[b]	
	4) WA5	

[a]This *Sma*I-RFLP type included isolates WA4, WA6, WA7, WA8, WA9, WA10 and WA13 ($n=7$).
[b]This type was indistinguishable from *Sma*I-M5, the predominant type found among 104 non-Aboriginal isolates.

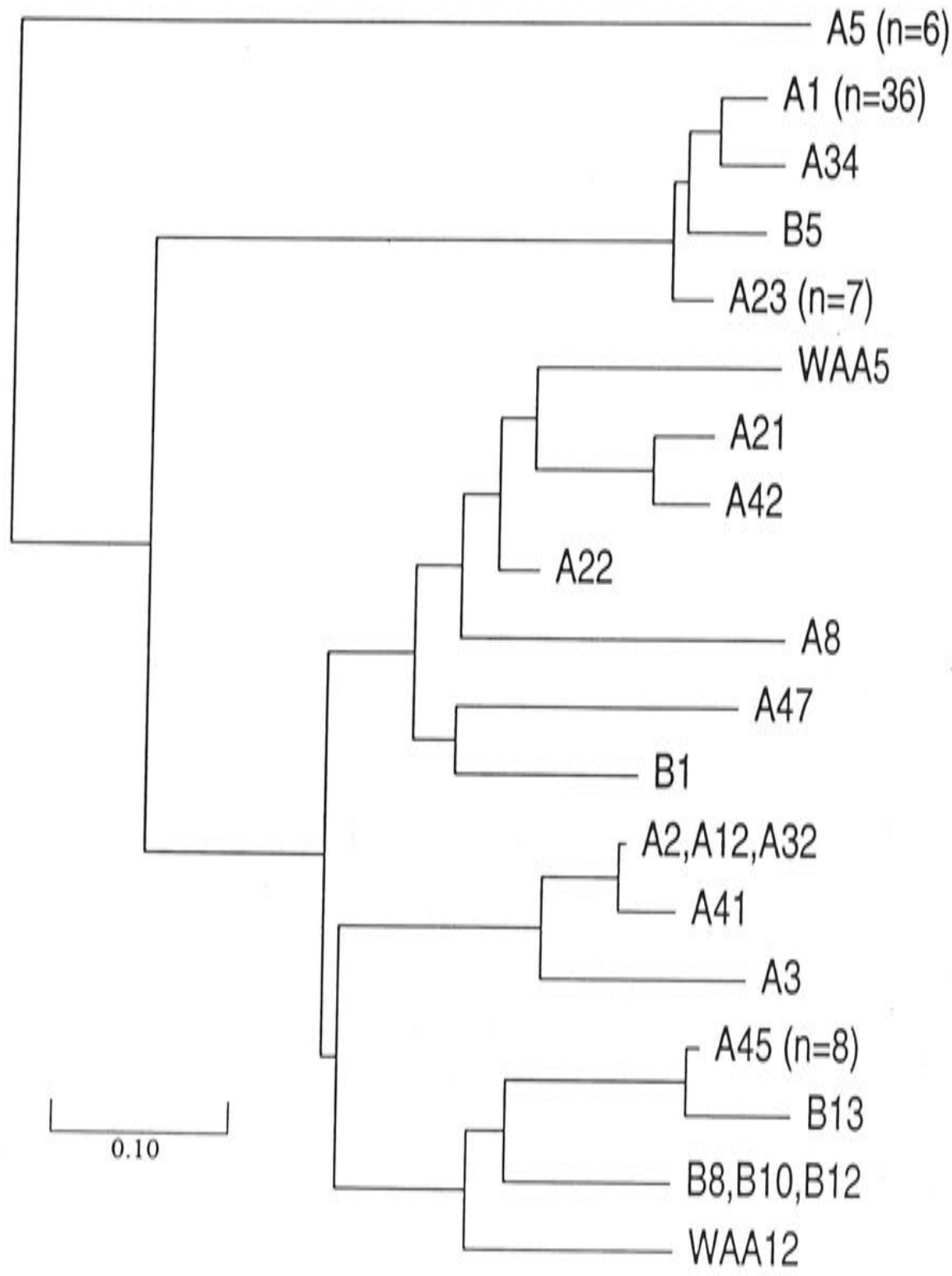


Figure 5.12 **Dendrogram showing the clustering of 19 *Sma*I-RFLP types found among 76 isolates obtained from Aboriginal patients in the Northern Territory and Western Australia.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbor-joining method and the TDRAW program was used to produce the appropriate tree. The bar indicates a genetic value of 0.10. The sum of the numerical value of the horizontal lines between pairs is an estimate of the genetic distance between the two types.

Types that fell in clusters ($n=3$) at $F \leq 0.5$	Types that fell in the clusters or branches ($n=14$) at $F \geq 0.9$		Pairs of types ($n=6$) at $F \geq 0.9$
1) A5	1) A1[a], A34	8) B1	A1-A23
2) A1, A34, B5, A23	2) B5, A23[b]	9) A2, A41	A1-A34
3) All the other types	3) WA5	10) A3	A1-B5
	4) A21, A42	11) A45[d], B13	A2-A41
	5) A22[c]	12) B8	A21-A42
	6) A8	13) WA12	A45-B13
	7) A47	14) A5[e]	

[a]This *Sma*I-RFLP type included isolates A1, A4, A6, A9, A11, A13, A14, A15, A17, A18, A20, A26, A27, A29, A31, A35, A36, A38, A39, A43, A46, A50, B6, B7, B11, B14, B15, B21, WA4, WA6, WA7, WA8, WA9, WA10 and WA13 ($n=36$).

[b]This *Sma*I-RFLP type included isolates A23, A28, A33, A44, A49, WA2 and WA3 ($n=7$).

[c]This type was indistinguishable from *Sma*I-M5, the predominant type found among 104 non-Aboriginal isolates.

[d]This *Sma*I-RFLP type included isolates A45, B2, B3, B4, B16, B18, B19 and B20 ($n=8$).

[e]This *Sma*I-RFLP type included isolates A5, A7, A16, A19, A37 and A45 ($n=6$).

5.3.3 Numerical index of the discriminatory ability of the typing system

Simpson's index of diversity (D) was 0.75 when applied to the data to determine the discriminatory power of *Sma*I among the Aboriginal isolates. No correcting factor for the small sample size (*n=76*) was applied in determining D but the poor discriminatory power seen here is due to the relatively few types found among the sample.

5.4 Discussion

Seventy-six isolates from Aboriginal patients and carriers living in three geographically distinct rural regions of Australia were obtained for this study. An outstanding feature of the genetic characterisation of these isolates was the predominance of a single type that was found in each geographical area. It was genetically distinct (*F*=0.34) from *Sma*I-M5 that predominated in the non-Aboriginal meningitis and epiglottitis sample and was designated *Sma*I-A1. Interestingly *Sma*I-A1 was indistinguishable from RFLP type *Sma*I-S37, that represented the single epiglottitis isolate in the Sydney sample and subsequently, provided evidence that this type can be associated with epiglottitis. Among the Aboriginal isolates *Sma*I-A1 was recovered from both healthy carriers and sick children including those with meningitis, pneumonia, gastroenteritis, cellulitis, bronchiolitis, and acute lower respiratory tract infections.

SmaI-A1 was the predominant type in the Alice Springs and Western Australian subsets but it was one of two types that appear to be endemic clones circulating in the Bathurst Island population. Even when the same types found among multiple isolates from the same child are counted as one member of a type the population structure of 2 endemic clones remains. A second major endemic type found among the Bathurst Island isolates was exclusive to this group except for one isolate from the Alice Springs region. It may represent a local clone that has spread to Central Australia.

The sample of Aboriginal isolates was further distinguished by the relatively large genetic distance between most types compared to that found among the non-Aboriginal isolates. The large genetic distances suggest that some types may have been long established in the communities in which they were found and are not the result of recent genetic changes. At least one member of a clonal pair is a type most likely to be due to recent genetic changes and there are few clonal pairs among the Aboriginal types.

Another distinguishing feature of the Aboriginal sample was its smaller ratio of types to isolates compared to that found among the 104 non-Aboriginal meningitis and epiglottitis isolates. Ratios of 0.25:1 and 0.45:1 were found among the Aboriginal isolates and the non-Aboriginal isolates, respectively. Thus, there were almost 50% more isolates per type in the Aboriginal sample. This observation raises questions that cannot be answered at this time. Does the isolation of the rural populations contribute to fewer endemic types? Or does the rate of transmission among Aboriginal populations result in a few endemic types? On the other hand, do the types contribute to increased transmission rates?

Small population groups among Aboriginals, overcrowded living conditions, poor hygiene and heavy colonisation rates may explain the presence of a few types exploiting an environment that supports a high rate of transmission. In contrast, non-Aboriginals who live in larger population groups but with less contact and lower colonisation rates are not as efficient in the transmission of Hib strains. This may result in the presence of more strains that are not transmitted as easily. It also suggests that the predominant non-Aboriginal clone may be more transmissible than others found in this population group. PFGE typing has provided a remarkable glimpse into the structure of Hib populations in these two epidemiologically distinct groups. Further study is needed to answer the questions it raises.

This study provided evidence that two genetic types of Hib can be present in individuals with Hib disease. An isolate, A22, that was recovered from the cerebrospinal fluid of a four month old child with meningitis from Hermmansburg, Northern Territory had an RFLP pattern that was indistinguishable from the predominant non-Aboriginal RFLP type, *Sma*I-M5. (Only one other isolate of this type was found and it was in the subset from metropolitan Perth.) A22 was also closely or clonally related to a second isolate, A21, that was recovered from the blood of this same child. There was a 4 fragment difference between the two types and they shared 24 of 28 fragments found between them.

Another pair of isolates, A34 and A35, recovered from the blood and cerebrospinal fluid, respectively, of a five month old child with meningitis were also found to be different types but were clonally related. A35, the isolate from the cerebrospinal fluid was the predominant Aboriginal type, *Sma*I-A1. In contrast each of two other pairs of

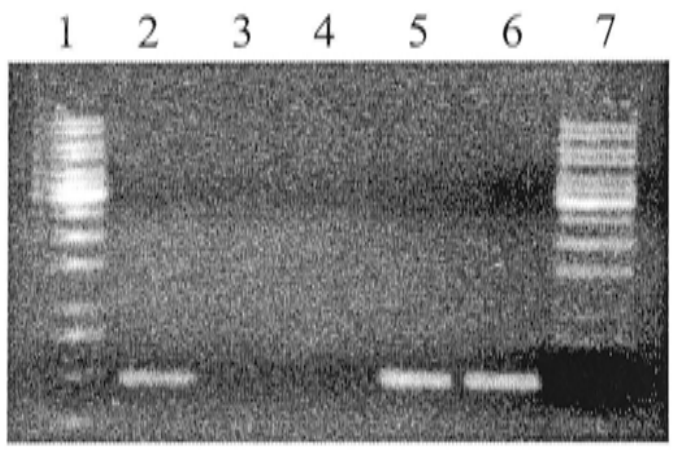
isolates from a child with pneumonia and a patient (age unknown) with meningitis and otitis media were types with RFLPs that were indistinguishable.

It has been established that a significant number (~35%) of invasive type b clinical isolates contain amplified capsule gene sequences and may possess three, four or even five copies of the *cap* B repeat (the capsule genes of Hib are arranged as a duplication of an 18 kb DNA segment (Hoiseth *et al.*, 1986)) and that *Sma*I does not cut within the 18 kb repeats but just outside the *cap* B locus (Corn *et al.*, 1993). The recombinational event leading to the amplification of *cap* B sequences should add DNA in 18-kb increments. Thus, it might be possible to detect capsule amplification when probing the *Sma*I fragments with pU082 if the cap locus was located on a fragment whose size increase could be detected by its decreased migration on the gel. Examination of the southern blot of *Sma*I fragments probed with pU082 showed that the *cap* locus was located on the largest fragment (~420 kb) of A22 and on two fragments (~420 kb and ~120 kb) of A21. This suggests that a genetic event may have resulted in the insertion or creation of a *Sma*I site within the *cap* locus but this needs further study to confirm.

The *cap* locus was found on the largest fragment of A35 (~620 kb) and the largest fragment of A34 (~640 kb) which are shown on Figure 5.4. Keeping in mind that estimation of fragment size using a plot of log molecular weight versus distance migrated is affected by the non-linear migration patterns of DNA found with PFGE, the difference in size between the fragments that hybridised was found to be approximately 18 kb. It seems possible that A35 represents an amplified state of A34. Another explanation is that an insertion occurred in another portion of the fragment that has nothing to do with the *cap* locus but I found the former explanation an intriguing speculation. Among others it raises the perhaps philosophical question as to whether or not an amplified state represents a different clone and thus, as one would interpret with this typing system (if fragment differences result between an amplified state and an unamplified state), a different genetic type. However, an attempt to demonstrate increased amplification product of the cap b locus in A35 was unsuccessful.

Note: To investigate if isolate A35 represents an amplified state of isolate A34 detection of type b specific capsular genes using PCR typing was performed as described in Chapter 2 in the methods section. Total chromosomal DNA concentration was estimated by comparing a small DNA sample with known

amounts of bacteriophage lambda DNA on an ethidium bromide stained electrophoretic gel and template DNA was carefully standardised. A35 was not observed to produce excess amplification product.



PCR amplification products:
Lanes 1 and 7 are 1 kb size marker
Lane 2 is Hib control
Lane 3 is NTHi control
Lane 4 is reaction control with no DNA template
Lane 5 is A34
Lane 6 is A35

It should be noted that one isolate provided for this study that was previously identified as Hib was shown to be a nontypeable *Haemophilus influenzae* because it did not hybridise with pU082. On examination, its RFLP pattern (Figure 5.3) was not similar to those found among the Hib isolates. This reminds us that false positives (and false negatives) are associated with conventional serotyping (Murphy and Apicella, 1987) and that capsular genotyping is necessary to confirm the presence of the *cap* locus in *Haemophilus influenzae* isolates. When *H. influenzae* serotyping is undertaken slide agglutination results should be interpreted within 30 seconds of mixing the reactants in order to avoid misclassification of non-encapsulated strains because of autoagglutination. This problem is exacerbated when the expectation is that capsular type b isolates are the most commonly recovered invasive *H. influenzae*. Occasional strains of *H. influenzae* may be non-typeable on initial isolation because of glycocalyx formation (Gratten and Montgomery, 1991). Glycocalyces are lost on serial subculture and antibody binding sites are re-exposed. Multiple subcultures, however, may encourage the loss of capsular material and hence type specificity.

In summary, the evaluation of 76 Aboriginal isolates using PFGE showed the population structure of Hib among Aboriginals is clonal and a single type predominates, not unlike the structure found in other studies of Hib populations. However, the predominant type is distinct from that found among non-Aboriginals in urban areas and overall the types are more diverse and fewer were found per isolate than among non-Aboriginals. These results follow in the footsteps of others. We first learned about a marked discrepancy between OMP subtypes causing Hib disease in the two major

groups of Australian children from Gilbert and Clements in 1993. This difference has now been confirmed using a genotypic typing system and these results show that two Hib types can be simultaneously present in a patient with Hib disease.

CHAPTER 6

Pulsed field gel electrophoresis used to investigate the genetic diversity of *Haemophilus influenzae* type b shows differences between Aboriginal and non-Aboriginal isolates

6.1 Introduction

The purpose of this chapter is to describe the genetic characteristics determined by combining all the Hib isolates studied previously in this project with additional isolates that have not formerly been evaluated. The additional isolates are from non-Aboriginal patients diagnosed with meningitis or invasive disease from the Alice Springs region and Townsville, Queensland and from patients with infections other than meningitis and epiglottitis from the Sydney region and Western Australia.

6.2 Methods

Preparation of DNA, restriction endonuclease digestion, PFGE conditions, determination of the numbers and mobilities of fragments, estimation of genetic diversity and construction of dendrograms and hybridisation were performed as previously described.

6.2.1 Selection of isolates for examination

One hundred-eighty isolates characterised by PFGE as described in Chapters 3-5 were included in the sample. Thirty-three non-Aboriginal isolates were added to them. Ten were isolates from non-Aboriginals living in the Alice Springs region, five of whom were diagnosed with meningitis and 5 with invasive disease. Two meningitis isolates were from Townsville, Queensland and a further 15 non-Aboriginal isolates from Sydney and 6 from Western Australia that were associated with diseases other than meningitis or epiglottitis, were included. The breakdown of these isolates according to disease association is shown in Table 2.1. Appendices A.1-A.6 describe all the isolates in more detail.

6.3 Results

All the additional isolates in this study hybridised with pU082 but no further splitting of types was detected by *cap* polymorphisms. All the non-Aboriginal isolates from the Alice Springs region ($n=12$) hybridised with the P2 probe and but no further splitting of types was detected.

6.3.1 Analysis of restriction fragment length polymorphisms

Non-Aboriginal isolates from the Alice Springs regions, Northern Territory (Figures 6.1, 5.1, and 6.3, Appendix A.5) Twelve isolates from non-Aboriginal patients with Hib disease were obtained from the Alice Springs region. They included 2 pairs of indistinguishable isolates recovered from the blood and cerebrospinal fluid of 2 patients. One of each of these was removed for the analysis. Eight types were found among the 10 remaining isolates including 1 that was indistinguishable from the predominant Aboriginal type and 2 that were indistinguishable from the predominant non-Aboriginal type. The predominant non-Aboriginal type was clonally related to one other type representing one isolate. The range of similarity among the types was 13-96%. Two isolates were indistinguishable from the most diverse type found in the Aboriginal subset that represented 5 isolates from the Alice Springs region. The distribution of isolates among types was:

Non-Aboriginal Alice Springs region isolates (*n=10*)

*Sma*I-RFLP type	Number of isolates/type
NT6[a], NT9[a]	2
NT8[b], NT11[b]	2
NT4[c]	1
NT12, NT1, NT10, NT2, NT7	1 each

[a]indistinguishable from the predominant non-Aborginal type
[b]indistinguishable from the most diverse Aborginal type
[c]indistinguishable from the predominant Aborginal type

Sydney isolates not associated with meningitis or epiglottitis (Figures 4.3-6 and 6.3, Appendix A.5) Fifteen isolates associated with cellulitis (n=2) and invasive disease not meningitis or epiglottitis (n=13) were obtained from the Sydney region. Eight *Sma*I-RFLP types were found among them. They shared 4 types with the meningitis/epiglottitis subset from Sydney and added 4 new types. Eight of the isolates were indistinguishable from the predominant non-Aboriginal type. The range of similarity among the types was 55-96%. The distribution of isolates among types was:

Sydney region isolates associated with cellulitis or invasive disease not meningitis or epiglottitis (*n=15*)

*Sma*I-RFLP type	Number of isolates/type
S1[a]	8
all others[b] (*n=7*)	1 each

[a]indistinguishable from the predominant non-Aborginal type; included S1, S3, S4, S23, S26, S31, S33, and S34
[b]included S10, S13, S15, S19, S21, S27, S32

Non-Aboriginal Western Australian isolates not associated with meningitis or epiglottitis (Figures 4.7-8, 6.3, Appendix A.4) Six isolates associated with Hib disease not meningitis or epiglottitis were obtained from Western Australia. Five were from metropolitan Perth and one was from a rural region. All the isolates except one from metropolitan Perth were indistinguishable from the predominant non-Aboriginal type.

Townsville isolates (Figures 6.2 and 6.3, Appendix A.5) Two isolates from patients with Hib meningitis each represented an RFLP type. One type was indistinguishable from the non-Aboriginal predominant type. The other type was distinct from it at an $F = 0.50$. Both of these isolates were collected after in 1993 sometime after July.

All the isolates combined *Note*: the names of RFLP types have now been assigned according to their position in the dendrogram that represents 213 isolates. How the types are named is explained in section 6.3.3 and is based on the major cluster and clonal grouping into which a type fell. What has been referred to as RFLP type *Sma*I-A1, the predominant Aboriginal type, is now known as type F2a and *Sma*I-M5, the predominant non-Aboriginal type, is now known as type A8b.

Well-resolved patterns of 13 to 17 *Sma*I fragments of approximately 8-650 kb representing 69 RFLP types were found among the 213 isolates. A total of 78 distinct fragments was observed on the gels. Only of these fragments (~43 kb corresponding to fragment 8.13 of HS008) was present in the RFLP patterns of all the isolates. It did not carry the *cap* locus and the P2 gene was found on a 43 kb fragment in only one of the 94 isolates typed with the P2 probe. Forty-eight (70%) of the RFLP types were represented by single isolates and 22 (32%) types had multiple isolates. The two most numerous types, one predominating in the Aboriginal subset ($n=37$) and one predominating in the non-Aboriginal subset ($n=55$), represented 43% of the sample. The remaining 19 types with multiple members were represented by two to eight isolates.

Nineteen types were found among the isolates from Aboriginal children ($n=76$) and 55 types were found among those from non-Aboriginal children and adults ($n=137$). The ratio of isolates to type was 4:1 and 2.5:1 among the Aboriginal and the non-Aboriginal populations, respectively, with only 5 types shared by both population groups. For the

most part, within each of the 5 shared types, the Aboriginal and non-Aboriginal isolates were from the same geographic region.

Type F2a accounted for 35 (46%) of the Aboriginal isolates but only 2 (1%) of non-Aboriginal isolates (Table 6.1). Type A8b accounted for 53 (39%) of the non-Aboriginal isolates and 2 (3%) of the Aboriginal isolates (Table 6.2).

Table 6. 1 Distribution of Hib *Sma*I-RFLP type F2a (n=37) that accounted for 46% of the Aboriginal isolates

Group and geographic location	Number of isolates									
	Meningitis		Epiglottitis		Other		Carrier		Total	
	A	n-A	A	n-A	A	n-A	A	n-A	A	n-A
Canberra region										
Victoria										
Sydney region				1						1
Perth, Western Australia	3								3	
Townsville, Queensland										
Alice Springs region	4	1			14[a]		4		22	1
Bathurst Island							6		6	
Rural Western Australia	4								4	
Total	11	1	0	1	14	0	10	0	35	2

[a]These isolates were associated with pneumonia (*n=5*), gastroenteritis (*n=3*), acute lower respiratory tract infections (*n=3*), cellulitis (*n=1*), bronchiolitis (*n=1*), and failure to thrive (*n=1*).

Table 6. 2 Distribution of Hib *Sma*I-RFLP type A8b (n=55) that accounted for 29% of the non-Aboriginal isolates

Group and geographic location	Number of isolates									
	Meningitis		Epiglottitis		Other		Carrier		Total	
	A	n-A	A	n-A	A	n-A	A	n-A	A	n-A
Canberra region		4		1						5
Victoria		4		7						11
Sydney region		11				6[a]				17
Perth, Western Australia	1	4		4		4			1	12
Townsville, Queensland		1								1
Alice Springs region	1					2			1	2
Bathurst Island										
Rural Western Australia		3		1		1				5
Total	2	27	0	13	0	13	0	0	2	53

[a]Inasive isolates (diagnosis unknown)

6.3.2 Genetic relationships determined by numerical analysis

F values ranged from 0.07 to 1, corresponding to a similarity range from 7-100%. An *F* value of 1, determined when all fragments between a pair were shared, corresponded to identical fragment patterns on the gels. Thus, when $F = 1$ the members of the pair are considered to be the same RFLP type. The smallest F value found, 0.07, corresponded to pairs where only 2 of the 14 and 15 fragments of the isolate pair were shared. A similarity matrix based on pairwise comparison of *F* for all RFLP types was used to calculate a dendrogram (Figure 6.3) as described in Methods.

The isolates fell into seven major clusters, designated Groups A-G, at an *F* value ≤ 0.5. There were 41 branches at an F value ≥ 0.9 corresponding to a genetic distance of greater than 90% similarity that are referred to as clonal groups. A single type or a cluster of closely related types whose members differ by 3 or fewer fragments represented each of the clonal groups. For comparative purposes each RFLP type was given a three character name according the major cluster (A, B, C, D, E, F, or G) in which it fell, the clonal group of which it was a member in that cluster (numbered from 1) and its unique type in the clonal group (lettered from a). Appendix C shows the distribution of types and isolates among the clusters and clonal groups found in dendrogram seen in Figure 6.3.

Two highly divergent groups (designated F and G) comprising 7 types represented 50 (66%) of the Aboriginal isolates and 6 (4%) of the non-Aboriginal isolates. A single genetically distinct group (designated A) comprising 39 types represented 7 (9%) of the Aboriginal isolates and 106 (77%) of the non-Aboriginal isolates.

The largest major cluster, group A, included 113 isolates and comprised 39 types in 15 clonal groups, 11 of which comprised multiple types. Type A8b that comprised 26% of the sample and 39% of the non-Aboriginal isolates fell in group A. Distribution of RFLP types among Groups B to E is shown in Figure 6.3. Group F had 48 members among 6 types in 4 clonal groups including type F2a that comprised 20% of the sample and 46% of the Aboriginal isolates. Group G included a single type, G1, comprised of 6 Aboriginal isolates and 2 Caucasian isolates from the Alice Springs region. G1 was the most highly divergent type in the sample with less than 15% similarity to either of the two predominant types, F2a and A8b ($F = 0.13$ and 0.14, respectively).

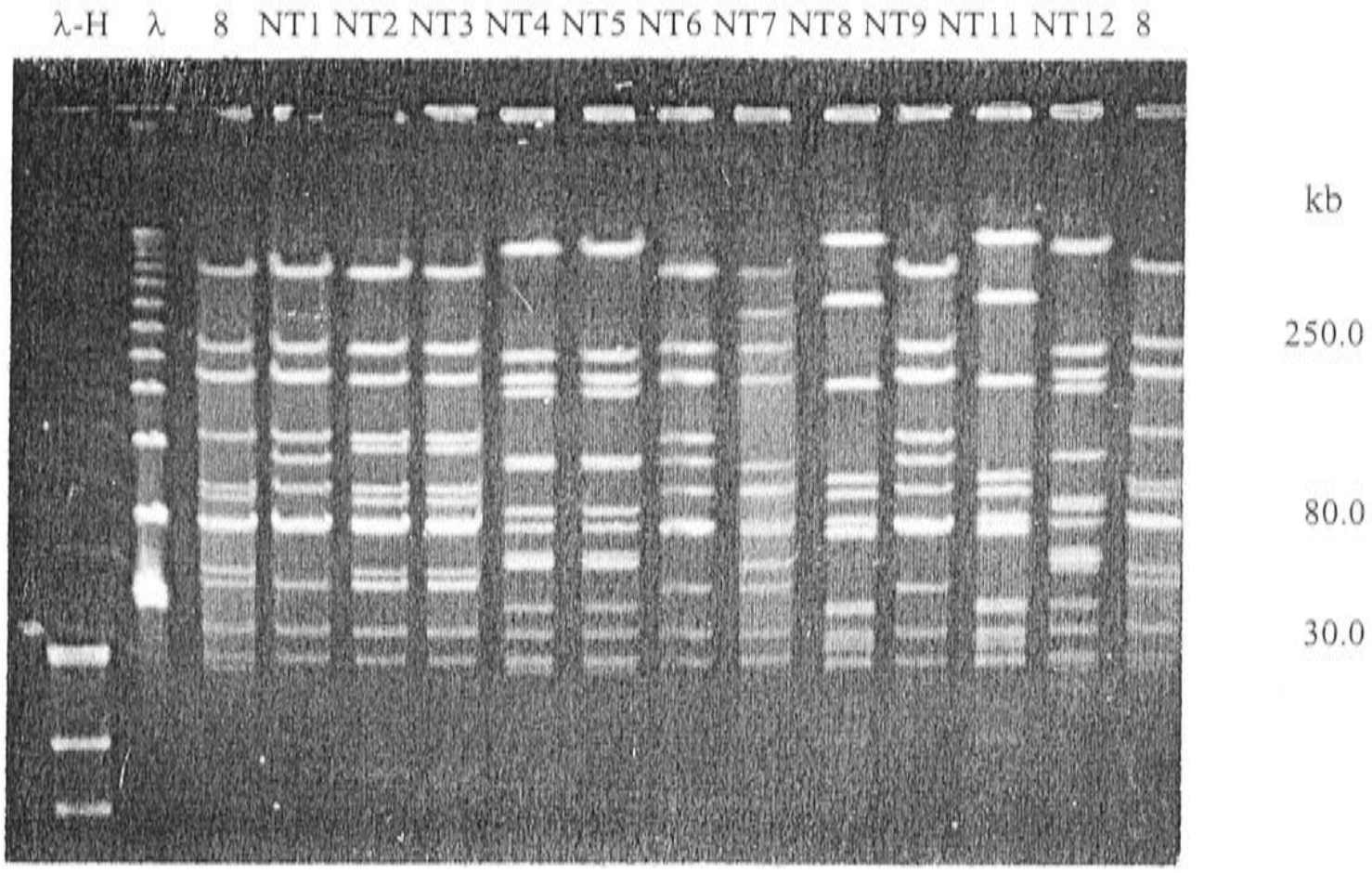


Figure 6.1 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 11 non-Aboriginal Hib isolates from the Alice Springs region. Lanes labelled 8, λ and λ-H contain fragment standards.

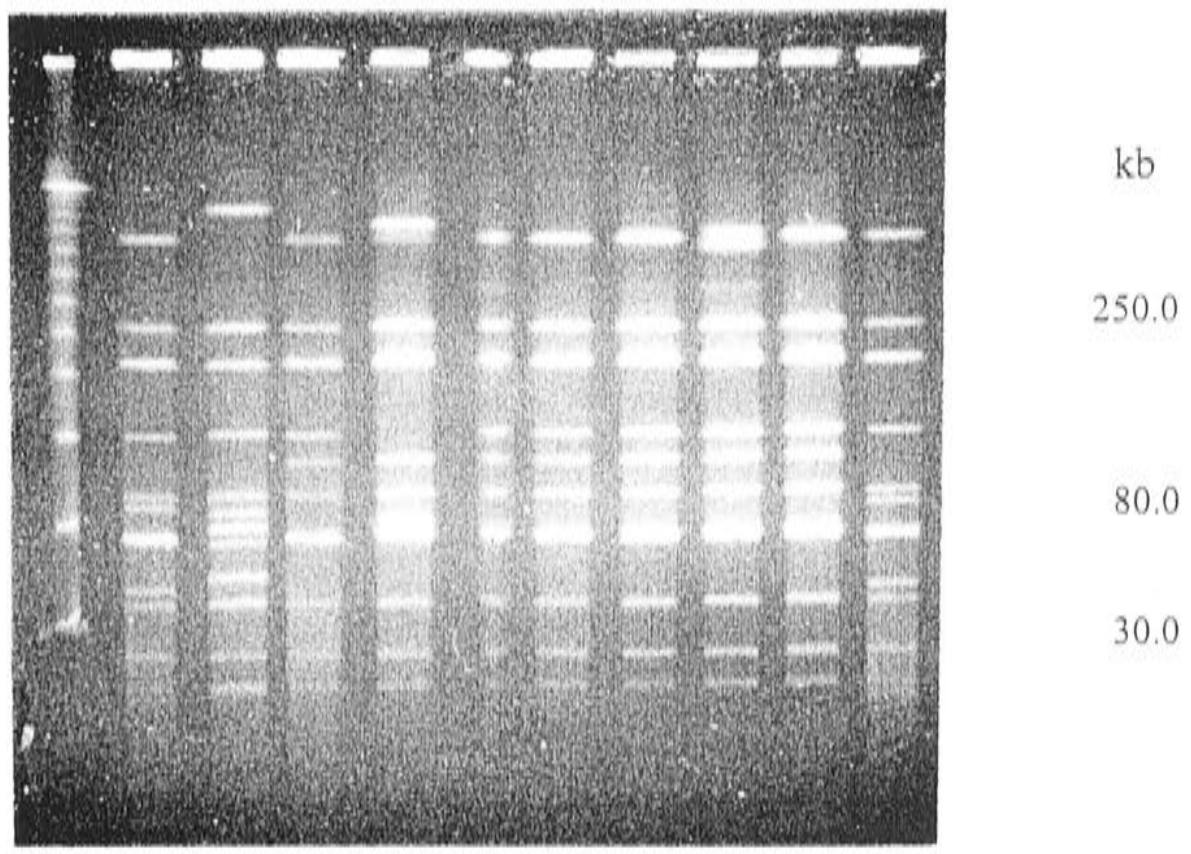
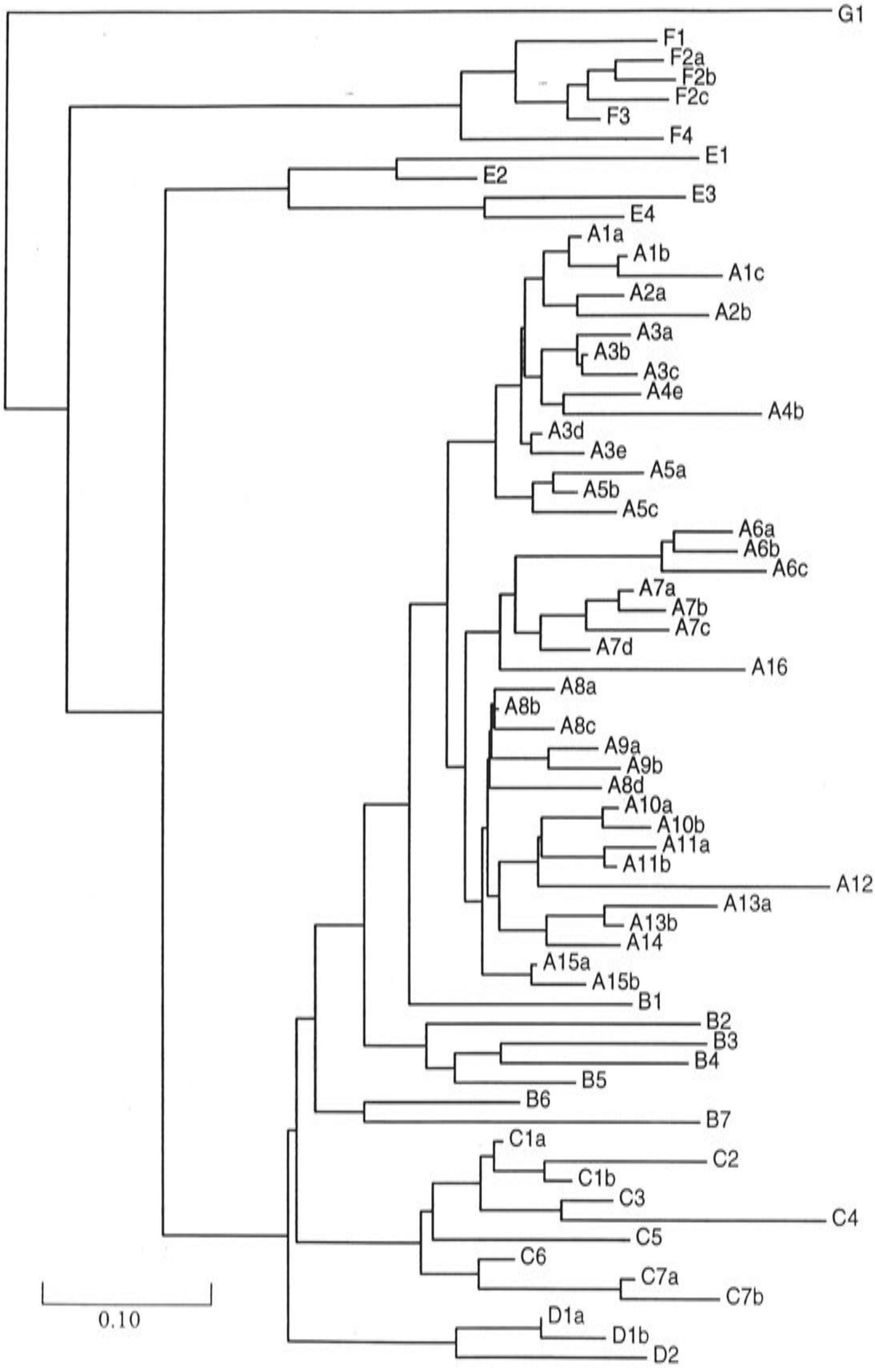


Figure 6.2 *Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 2 Hib isolates from meningitis patients from Townsville, 3 from Sydney, and 3 from Canberra. Lanes labelled 8 and λ contain fragment standards.



Group	Aboriginal	Non-Aboriginal	Total
G1	6	2	8
F1	0	1	1
F2a	**35**	**2**	**37**
F2b	1	0	1
F2c	1	0	1
F3	7	0	7
F4	0	1	1
E1	0	1	1
E2	0	2	2
E3	0	1	1
E4	0	1	1
A1a	0	4	4
A1b	0	2	2
A1c	0	1	1
A2a	0	1	1
A2b	0	1	1
A3a	0	1	1
A3b	0	1	1
A3c	0	1	1
A4a	0	1	1
A4b	1	1	2
A3d	0	1	1
A3e	0	1	1
A5a	0	1	1
A5b	0	1	1
A5c	0	1	1
A6a	0	1	1
A6b	0	1	1
A6c	0	1	1
A7a	0	2	2
A7b	1	0	1
A7c	1	0	1
A7d	0	1	1
A16	1	0	1
A8a	0	2	2
A8b	**2**	**53**	**55**
A8c	0	1	1
A8d	0	1	1
A9a	0	2	2
A9b	0	3	3
A10a	0	1	1
A10b	0	1	1
A11a	0	1	1
A11b	0	8	8
A12	1	0	1
A13a	0	1	1
A13b	0	2	2
A14	0	1	1
A15a	0	3	3
A15b	0	1	1
B1	1	0	1
B2	0	1	1
B3	0	1	1
B4	0	1	1
B5	0	1	1
B6	0	1	1
B7	0	3	3
C1a	0	3	3
C2	0	1	1
C1b	0	4	4
C3	1	2	3
C4	0	1	1
C5	3	0	3
C6	0	1	1
C7a	8	0	8
C7b	1	0	1
D1a	3	0	3
D1b	1	0	1
D2	1	0	1

Figure 6.3 Dendrogram showing the clustering of 69 *Sma*I RFLP types of 213 Hib isolates obtained from Aboriginals and non-Aboriginals from rural and urban regions around Australia. The dendrogram was generated using the mathematical model of Nei and Li and distances were calculated by the neighbor-joining method. The distance between any two taxa is the sum of the horizontal lines between them. The bar indicates an F value of 0.10 or 10% similarity. The isolates fall into seven major clusters designated A through G. There are 41 branches or clusters at $F \geq 0.90$. Only types G1, F2a, A4b, A8b, and C3 share both Aboriginal and non-Aboriginal members.

The largest genetic distance estimate was between two members of Group A (types A2b and A10) and the members of Group G. They were separated by genetic distances of $F = 0.07$ with only 2 of 29 fragments shared between pairs.

6.3.3 Genetic diversity and correlation to epidemiology

Clustering of RFLP types was not associated with disease type. Isolates from patients with meningitis and epiglottitis, as expected, were each represented by a diverse range of types and no single type was associated with a particular diagnosis or carriage. In addition, no association was found between age, sex, date of collection or geographic location. A strong association, however, between both the origin of isolates from Aboriginal children and RFLP type F2a and the origin of isolates from non-Aboriginal children and RFLP type A8b (chi square, $p < 0.001$) was found and 3 other RFLP types (F3, G1, and C7a) with multiple members were composed of predominantly Aboriginal isolates.

Type F3 comprised 7 isolates and fell in the same clonal group as type F2a. It accounted for 5 isolates from Aboriginal children living in the Alice Springs region and 2 isolates from Aboriginal children living in rural Western Australia. Type G1 included 6 Aboriginal isolates and 2 non-Aboriginal isolates from the Alice Springs region. Eight isolates from Aboriginal children shared RFLP type C7a that fell in a group comprised of 13 Aboriginal isolates and 11 non-Aboriginal isolates. The association of type C7a with Aboriginality or geographic location is not statistically significant, however, 7 of the 8 isolates in type C7a were from Bathurst Island in the Northern Territory, as was the single member in a clonally related type, C7b.

6.3.4 Numerical index of the discriminatory ability of the typing system

Simpson's index of diversity (D) was 0.90 when applied to the data to determine the discriminatory power of *Sma*I among the entire sample of 213 Hib isolates.

6.4 Discussion

The sample of 213 Hib isolates comprised 69 different types of *Sma*I-RFLP patterns with 17 or fewer fragments that were well resolved by PFGE. Closely related and more distantly related genomic types were easily discriminated.

When Simpson's index of diversity was applied to the system an index of D = 0.90 was calculated. This indicates that the system demonstrates good discriminatory power for typing based on the probability of two unrelated strains being characterised as the same type. An index greater than 0.90 is considered desirable if the typing results are to be interpreted with confidence but assessment of other factors such as typability and reproducibility affect the usefulness of a typing system. Cost, level of technical skill required and time requirements will also affect whether a system is used or not.

The *Sma*I RFLPs fell into 7 major clusters in a tree-like dendrogram that is consistent with a clonal population structure, as is the presence of 2 predominant types among the 69 types found. The two predominant types, F2a and A8b, accounted for 43% of the sample and included almost one-half of the Aboriginal isolates and well over one-third of the non-Aboriginal isolates, respectively. Not surprisingly, the results described a heterogenous clonal population structure and extend the geographical range over which the clonality of Hib has been demonstrated.

Differences and similarities between Hib isolates were readily visualised from the CHEF patterns and a general grouping of the isolates could be established in this way. However, a more quantitative approach can be derived by subjecting the data to the analytical procedures described by Nei and Li. In an attempt to describe what are otherwise qualitative differences in fragment patterns among a large body of comparative data, we estimated DNA similarities on the basis of the fraction of common fragments generated by the endonuclease. The value generated is referred to as the F value.

It has been suggested by Tenover that, in the context of analysing outbreak strains using one endonuclease, a 1-3 fragment difference can occur between closely related strains and possibly related strains can be found among those that differ by 4-6 fragments (Tenover *et al.*, 1995). It has also been reported that closely related strains of *E. faecalis* can differ by up to six bands and that patterns have been shown to differ by up to seven bands and still be clonally related (Thal *et al.*, 1997). Furthermore, using hybridization analysis and PFGE in the investigation of insertion mutations in the transferrin binding systems of Hib, Curran et al. found 5 different changes of 0-3 bands between *Sma*I fragments of 5 insertion mutants. These results remind us that several different chromosomal mutations could generate the same phenotype.

Thus, the values of F obtained should not be seen as precise estimates of genetic distance but summary values indicative of overall similarities and differences between isolates and a means to illustrate relationships within a large body of comparative data. F values $\geq$0.9 representing pairs of isolates that differ by up to 3 fragments were used to determine closely related isolates in this study and we describe such closely related isolates as a clonal group. However, pairs of isolates that are closely related can be found as members of different clonal groups as well as within clonal groups so the number of branches is not an absolute representation of all closely related pairs. For example, the predominant RFLP type, A8b, is found within a clonal group comprised of 2 other closely related types. It is also closely related ($F \geq 0.9$) to 16 types found in 10 other clonal groups within Group A.

Despite the wide diversity of the strains studied, including clinical and non-clinical isolates, it was not possible to detect an obvious correlation between a given genotype and the specific disease it can cause or a healthy state of the host. After subjecting the data to statistical analyses (averages distances and chi square analysis) no association between RFLP type and age, sex, date of collection, or disease manifestation was revealed. Because of the lack of Aboriginal isolates from urban areas and non-Aboriginal isolates from rural areas the association between geographic location and type was not found to be significant in this sample. The genetic distance separating the important major lineages ranged from a similarity of 20% to 50% yet there is apparently an equivalent ability to cause disease among the lineages. One of the patterns of Hib disease in Australia was a significant difference in the incidence of epiglottitis among urban populations. The incidence of epiglottitis in Victoria (Gilbert *et al.*, 1990) and the Australian Capital Territory (McGregor *et al.*, 1992) was twice that of Sydney (McIntyre *et al.*, 1991) and Western Australia (Hanna *et al.*, 1992). Another pattern of Hib disease in Australia was the lack of epiglottitis among Aboriginal children, similar to the epidemiology of Hib disease among other high-risk indigenous populations (Hanna, 1992). The possibility that differences in the incidence of epiglottitis among different population groups are related to differences in the strains of Hib has not been demonstrated in this study and the patterns of the incidence of epiglottitis in Australia remain unexplained. Other restriction enzymes which sample different areas of the genome may detect molecular differences that demonstrate clonal disease specificity but we did not detect such a relationship among the *Sma*I RFLP types in this sample.

Analysis of isolates from Aboriginal children living in geographically isolated regions of Australia revealed a relatively small number of genetically distinct RFLP types. A strong association was found between the origin of isolates from Aboriginal children and one of these types. Isolates from non-Aboriginal populations were more diverse and genetically distinct from most Aboriginal isolates; and, one of these types found in a cluster of closely related isolates was strongly associated with the origin of isolates from non-Aboriginal children. The data support the hypothesis described by Musser *et al.* (1990) that a causal relationship exists between the degree of ethnic mixing of human populations and the degree of diversity in clonal composition of Hib populations. According to this hypothesis, a large component of the current geographic variation in clonal composition of Hib reflects an older pattern of differentiation that evolved in relative geographic isolation before the Age of Exploration (beginning about 450 years ago) and has not been completely obscured by recent demographic changes (Musser *et al.*, 1990). It explicitly predicts that isolates from Aboriginal populations largely belong to a distinctive set of clones, as described in this study.

If Hib had been present in Aboriginal populations before European settlement, the striking difference between the predominant Aboriginal type and the predominant non-Aboriginal type suggests separate evolution of strains. The time span since permanent European settlement began on the East coast of Australia in 1788 does not support a long history of differentiation that may be needed for such a deep divergence between types. Therefore, it is more likely that urban non-Aboriginal types have been recently introduced from other areas of the world rather than derived from the Aboriginal types. Nonetheless, more extensive analysis is needed to provide data to support the explanation of the phylogenetic relationship between the Hib types described here.

Prior to the introduction of vaccination, the incidence of Hib disease in Aboriginal populations was reported to be extremely high and varied among communities. The estimated annual incidence among Aboriginal children under five in Central Australia (an area corresponding to the Alice Springs region in this study) was 991 cases per 100,000, with a high proportion of cases of meningitis and a case-fatality rate of 8.3% (Hanna, 1990). The corresponding rate among non-Aboriginal children in Central Australia was 215 per 100,000 which was significantly higher than that in other areas of the Northern Territory and four times higher than that in Melbourne (Hanna, 1992). It was concluded that environmental factors and cultural factors were responsible for the

very high incidence in Aboriginal communities (Hanna, 1990). However, this does not explain the increased incidence in non-Aboriginal children. An alternative explanation, that the strains circulating among Aboriginal children and affecting some non-Aboriginal children in the same geographic area are more virulent than urban strains has not been investigated.

Our data show that a few distinct types are responsible for Hib disease among Aboriginal populations; and, the high incidence of disease in very young children indicates very high transmission rates associated with environmental factors. According to Ewald's cultural vector hypothesis increased transmission rates and repeated passage through human hosts may lead to selection of more rapidly multiplying and virulent strains (Ewald, 1994). If so, such strains of Hib in Aboriginal populations are also likely to spread to non-Aboriginal children in the same geographic area and result in an increased diversity of strains in this group as well as an increased incidence of disease.

Differences found between Aboriginal and non-Aboriginal Hib types using PFGE demonstrate its practical discriminatory power for analysing large numbers of Hib isolates. Analysis of the huge similarity matrices required for large samples has been facilitated by the availability of computer programs and, with standardisation of PFGE methods, would allow inter- and intra-laboratory comparison of results in a common database. This would improve surveillance by increasing the capacity to detect new strains of Hib and monitor the presence of old strains. The usefulness of the system was shown when 2 isolates of Hib recently recovered in a nursing home outbreak in the Sydney region (Heath *et al.*, 1997) were analysed, and both were shown to have patterns identical to that of *Sma*I type A8b. This suggested that these isolates were not a new genotype since they were identical to the predominant type found in Sydney in the pre-vaccine era. This system will be useful for monitoring future Hib isolates from children in whom the vaccine has failed, unimmunised children, and adults. Further, the genotyping of clinically significant and carrier Hib isolates may help to elucidate transmission routes in community infections, endemicity, and the reasons for vaccine failures.

CHAPTER 7

ERIC-PCR and OMP26 Hybridisation Studies

7.1 Introduction

The development and extensive use of high resolution molecular typing systems based on direct analysis of genomic polymorphisms have greatly improved the understanding of the epidemiology of infectious disease. These systems exploit genetic variation and analyse DNA variation. As all genetic differences between individual bacterial cells are related to the primary sequence of their genomic DNA, the most direct method of identifying an individual isolate would be to determine this sequence for the whole genome. This is currently not feasible. The panoply of molecular typing methods that are in use vary in their ability to provide the answers to the questions sought and comparative studies using a combination of systems is often recommended (Struelens, 1998) (van Belkum *et al.*, 2001).

PFGE is now the "gold standard" for epidemiological typing and calculation of restriction pattern similarity coefficients and graphical display of pattern relatedness as dendrograms is also useful for interpretation, particularly for large scale studies (Struelens, 1998). Although this analysis has been criticised as invalid for phylogenetic inferences because DNA restriction fragment pattern variation is not due to independent events, it is supported by population analysis, e.g., of *Pseudomonas aeruginosa* (Struelens, 1998). For Hib, the DNA restriction polymorphisms detected by PFGE reflected evolutionary divergence with that identified by MLEE (Arbeit *et al.*, 1990). Its discriminatory power has been shown to be better than MLEE, ribotyping, and RAPD-PCR methods and is best used to identify the micro-variation that is required to distinguish between strains; it is suitable for typing both highly clonal and less clonal populations (Struelens, 1998). However, like other electrophoretic methods, it is not known what each fragment represents, i.e., the genetic basis of differences in band patterns is not known. To better understand the usefulness of interpreting PFGE fragment patterns for phylogenetic relationships characterisation of the fragments is warranted, as well as, comparison with other genetic methods.

To comparatively evaluate the data provided by PFGE two other molecular tests were applied to a selection of isolates from this study.

ERIC Enterobacterial repetitive intergenic consensus (ERIC) sequences are conserved regions of DNA dispersed throughout the genomes of Gram-negative enteric bacteria. ERIC sequences are highly conserved at the nucleotide sequence level, but their

chromosomal locations differ between species or strains (Gomez-de-Leon *et al.*, 2000). These elements have been successfully used to differentiate strains of *Bartonella*, *Bacillus subtilis*, *Citrobacter diversus*, methicillin resistant *Staphylococcus aureus*, *Streptococcus pneumoniae*, *Neisseria gonorrhoeae* and *Neisseria meningitidis* and other bacterial genomes (Olive and Bean, 1999). And, ERIC-specific primers have been used to produce genetic fingerprints *H. influenzae* strains (van Belkum *et al.*, 1994; Smith-Vaughan *et al.*, 1998; Gomez-de-Leon *et al.*, 2000; Pettigrew *et al.*, 2002). The largest practical application of ERIC-PCR to capsulated and unencapsulated strains of *H. influenzae* demonstrated the ability of this method to detect genetic diversity when 69 fingerprint patterns were found among 92 isolates recovered from Mexican children (Gomez-de-Leon *et al.*, 2000). Intriguingly, the results of this study show that some of the strains grouped according to the type of infection (i.e., meningitis strains versus other disease isolates). Further study is required to determine whether meningitis isolates present specific characteristics related to ERIC motifs and if these, in turn, may be recognised as having some involvement in the pathogenesis of Hib meningitis strains (Gomez-de-Leon *et al.*, 2000). In the present study genomic DNA amplified by PCR with primers for the ERIC sequences* was used to genotype Hib isolates and the discriminatory power of ERIC-PCR was compared with PFGE.

*True ERIC sequences are not present in the fully sequenced *H. influenzae* Rd so it is likely that ERIC fingerprints reflect random hybridisation of the primers (Pettigrew *et al.*, 2002).

OMP26 OMP26 is a major outer membrane protein of approximately 26-kDa isolated and purified from NTHi strain 289. It has been investigated for its suitability as a vaccine candidate against NTHi infection for which it shows significant potential (Kyd and Cripps, 1998; El-Adhami *et al.*, 1999). DNA sequences encoding OMP26 were previously cloned by PCR amplification using oligonucleotide primers designed from the equivalent gene of *H. influenzae* Rd (El-Adhami *et al.*, 1999). The PCR product of about 618 bp was digested and ligated into pQE-30 to give pP26/1. Ligated DNA transformed into *E. coli* XL-1 Blue was kindly provided by Dr Adam Smith, Canberra Centre for Mucosal Immunology.

When the pP26/1 construct was originally prepared hybridisation analysis of 19 different isolates of NTHi and a capsule deficient type b strain was performed. Genomic

DNA of all strains hybridised with the probe indicating similar OMP26 genes existed in all these strains (El-Adhami *et al.*, 1999). A BLAST® search for sequences deposited in public sequence databases that produced significant alignments with that of the NTHI–289 OMP26 gene (accession no. AF109085) returned 24 matches (at an effective value of 0) of previously characterised genes from *H. influenzae* isolates. Homology ranged from 96% to 99%. No other significant alignments were returned.

In this study a probe prepared from p26/1 was used to detect the location of the OMP26 gene among the DNA fragments produced by PFGE of Hib DNA digested by the restriction endonuclease *Sma*I. The purpose of this was to confirm that fragments with similar migration profiles on the electrophoretic gels carry similar genetic information.

7.2 Methods

7.2.1 Bacterial isolates

ERIC-PCR Twenty-nine isolates were selected from the collection of Hib strains that had been previously characterised by PFGE. Six isolates from each of the two predominant *Sma*I-RFLP types, F2a and A8b, were included in the selection. The other 17 isolates represented all the major groups delineated by the *Sma*I PFGE and types closely related to the predominant types as listed in Table 7.1. The number of RFLP types among the sample was 19. Isolate information can be found in the appendices.

OMP26 hybridisation Forty-nine isolates from the collection of Hib strains that had been previously characterised by PFGE were analysed for the presence of the OMP26 gene. Southern blots prepared from the electrophoretic gels seen in Figures 4.2, 4.3, 4.5, 4.8, and 5.8 that had been previously used for *cap* and P2 hybridisation studies were used in this study. The selection included 23 and 6 isolates, respectively, from the predominant *Sma*I-RFLP types, A8b and F2a. One other type had multiple isolates (n=4) and 15 types were represented by one isolate. The isolates are listed in Table 7.2. Four NTHi including NTHi-289, also previously characterised by PFGE and probed for the *cap* locus, were used as controls.

7.2.2 ERIC-PCR

DNA preparation Total chromosomal DNA was harvested using the Genomic Tip 20/G kit (Qiagen) from overnight sBHI broth cultures inoculated with a single colony. The DNA concentration was estimated by comparing a small DNA sample with known amounts of bacteriophage lambda DNA on an ethidium bromide stained electrophoretic gel. The DNA was dissolved at a concentration of ~20 ng/µl in TE.

ERIC-PCR and amplification conditions PCR and amplifications conditions are described in Chapter 2. The results were analysed only if control samples without extraneous DNA failed to yield amplified product. To check for reproducibility and to optimise performance 3 strains were typed on three or more separate occasions from independent DNA preparations. All isolates were fingerprinted at least twice.

ERIC-PCR patterns and analysis The numbers and size of fragments were determined by visual examination of the digitised photographs of the stained gels. Because template concentration and run parameters affect product, bands that were present on two analyses were included. The range of similarity or genetic distance, identification of clonal and the most diverse relationships among types and construction of a dendrogram was determined from a matrix produced using the mathematical model of Nei and Li as previously described for PFGE. A matrix of *F* values for all pairs of isolates was constructed and a dendrogram computed using the NJTREE program and the companion TDRAW program.

7.2.3 OMP26 typing

Southern blotting and hybridisation *Sma*I digested DNA fragments separated by PFGE were transferred onto nylon membranes as described in Chapter 2. Membranes were stored at 4°C until used. Extraction of purified plasmid pQE-30/OMP26 was accomplished using the Quantum Prep® Plasmid Miniprep Kit (Biorad). A digoxigenin-labeled probe was prepared, hybridised to the target filter under standard conditions, and the probe-positive bands were visualised by an immunochemiluminescent detection system according to the manufacturer's instructions (DIG-High Prime DNA Labelling and Detection Kit, Boehringer Mannheim).

7.3 Results

General characteristics and diversity of ERIC-PCR types ERIC-PCR produced banding patterns of 6-10 fragments that were not too complex to analyse. Two invariant fragments were found in each pattern. Another fragment was found in all the samples except NT8. Seven ERIC types (designated I-VII) were found among the 29 isolates. Two types accounted for 23 of the closely related isolates and the other 5 types represented each of the more diverse isolates. Table 7.1 shows the comparison of ERIC-types with PFGE groups, the source and the diagnosis of each isolate. The major divisions of Hib among the predominant PFGE types was found among the ERIC types though different PFGE types had the same ERIC type. Each of the A8b PFGE types and clonally or closely related types fell into ERIC-III. Each of the F2a types and their related types fell into ERIC-II, a type that is distinctly different from ERIC-III. The other isolates representing PFGE types selected from genetically diverse groups fell into different ERIC types.

Relationships among ERIC-types compared to PFGE groups The distance matrix constructed using ERIC fragment profiles is in table 7.2 and the dendrogram constructed from it is in Figure 7.1. The polymorphism obtained with ERIC-PCR allowed the 29 Hib isolates to be distributed in to a dendrogram into 2 groups at a genetic distance of $F \leq 0.50$; the isolate NT8 (PFGE type G1 and ERIC type I) was in one group and all the other isolates fell into the other group. The larger cluster formed 4 smaller clusters including one in which included 2 types related to PFGE-A8b/ERIC-III at a genetic distance of $F \geq 0.90$. The clustering of isolates was in agreement with that found among PFGE types (see Figure 6.3) although the genetic distance between the two predominant types, F2a and A8b, is not as distinct among the ERIC-PCR types as among the PFGE types. The F value determined using PFGE RFLPS between M15 (an isolate representing ERIC type II) and M5 (an isolate representing ERIC type III) is 0.4687 (see Table 3.11 in Appendix B) while the F value between ERIC type II and III is 0.6250 indicating greater similarity. When the distance matrix of a representative PFGE isolate for each ERIC type was calculated and compared to the ERIC type distance matrix by the program Diplomo a correlation of 0.90 was obtained. Simpson's index of diversity was calculated to be D = 0.89.

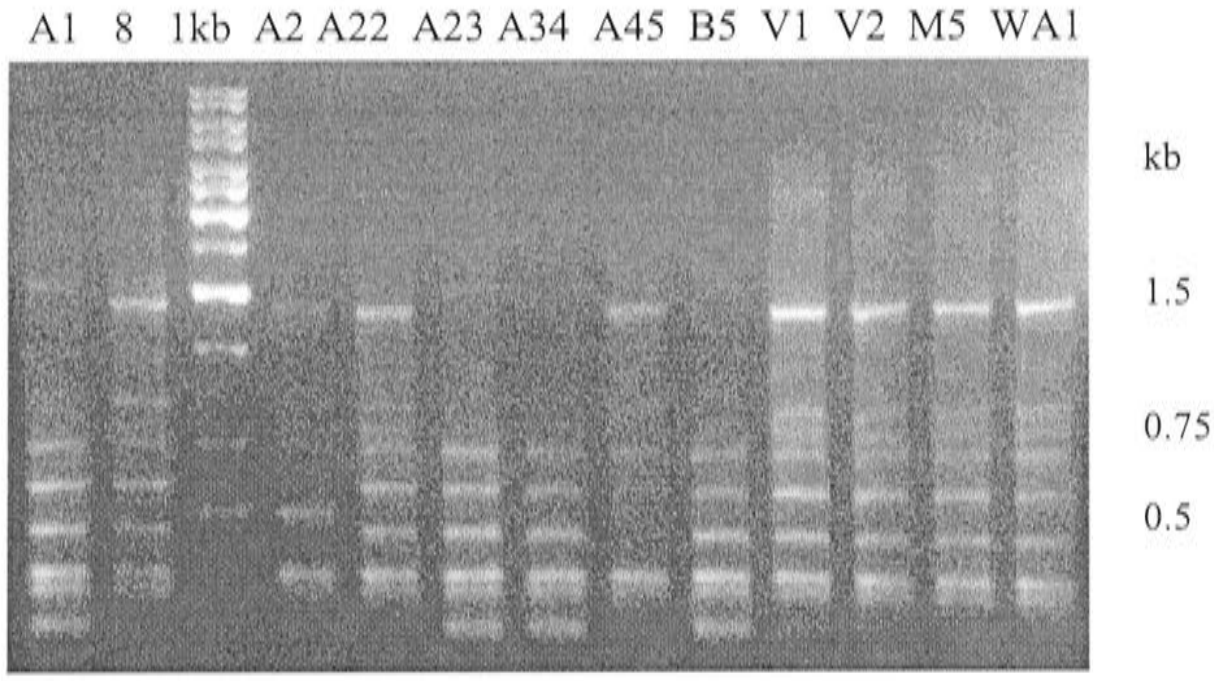


Figure 7. 1 ERIC-PCR fingerprinting patterns of Hib

The image was digitally acquired by using Bio-Rad GelDoc 1000 and Microsoft Powerpoint and annotated with Corel Draw and Microsoft Word. A size standard (GeneRuler™ 1kb DNA ladder) is in the lane marked 1kb.

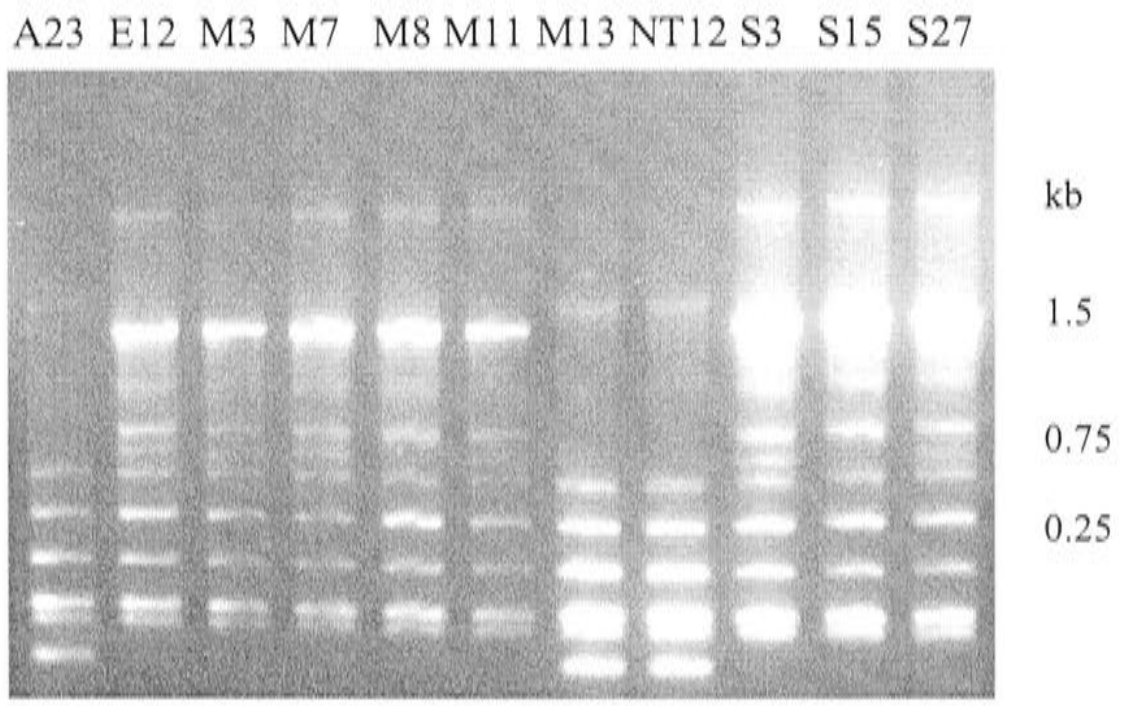


Figure 7. 2 ERIC-PCR fingerprinting patterns of Hib

The image was digitally acquired by using Bio-Rad GelDoc 1000 and Microsoft Powerpoint and annotated with Corel Draw and Microsoft Word.

Table 7. 1 Distance matrix of the *F* values of 7 ERIC-PCR types

IV	0.3750					
V	0.4286	0.8889				
VI	0.5000	0.7500	0.8571			
VII	0.5000	0.7000	0.8888	0.7500		
II	0.3077	0.5882	0.7777	0.6158	0.4706	
III	0.4000	0.9474	0.9412	0.8000	0.7368	0.6250
	I	IV	V	VI	VII	II

Table 7. 2 Comparison of ERIC-PCR types with PFGE typing groups

Isolate	Source	Diagnosis	PFGE Group	ERIC-PCR Type
NT8	bld	invasive	G1	I
M13	bld/csf	meningitis	F1	II
A1	bld	meningitis	F2a	II
A50	nk	gastro	F2a	II
B6	naso	carrier	F2a	II
NT4	bld/csf	meningitis	F2a	II
S37	bld	epiglottitis	F2a	II
WA4	csf	meningitis	F2a	II
A34	bld/csf	meningitis	F2b	II
B5	naso	carrier	F2c	II
A23	bld	febrile	F3	II
NT12	bld	invasive	F4	II
M7	bld/csf	meningitis	E1	III
V1	nk	meningitis	E2	III
M3	bld/csf	meningitis	A1a	III
NT2	bld	meningitis	A4b	III
WA42	bld	epiglottitis	A8a	III
A22	csf	meningitis	A8b	III
M5	csf	meningitis	A8b	III
WA1	csf	meningitis	A8b	III
V2	nk	meningitis	A8b	III
E12	bld	epiglottitis	A8b	III
S3	bld	invasive	A8b	III
S27	bld	invasive	A8c	III
M11	bld/csf	meningitis	A8d	III
M8	bld/csf	meningitis	A9a	IV
S15	bld	invasive	B7	V
A45	nk	pneumo	C7a	VI
A2	bld	meningitis	D1a	VII

Abbreviations: bld-blood, csf-cerebrospinal fluid, naso-nasopharynx, nk- not known

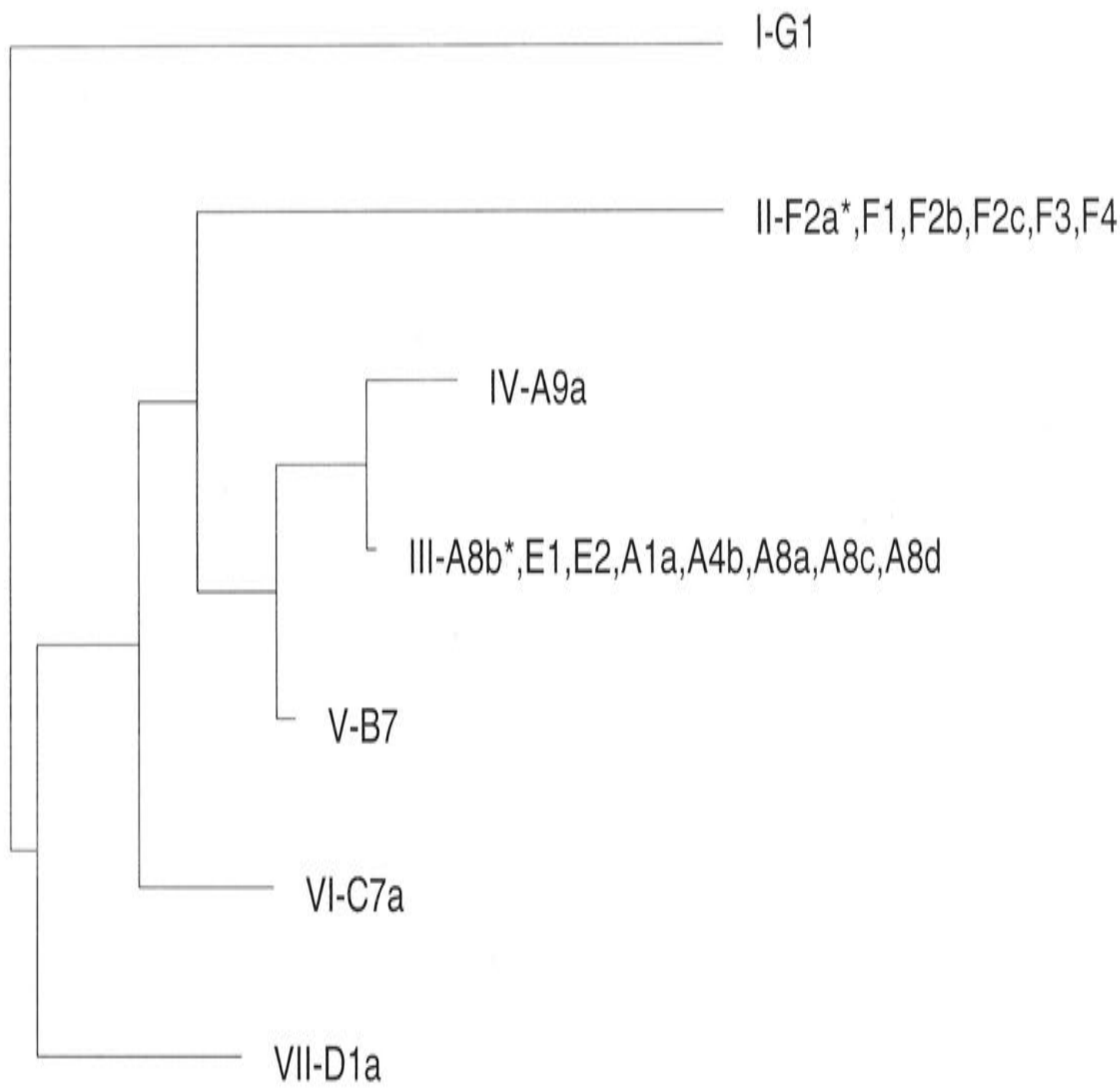


Figure 7. 3 **Dendrogram showing the clustering of 7 ERIC types found among 29 isolates selected from the major PFGE groupings.** The dendrogram was generated using the mathematical model of Nei and Li. Distances were calculated by the neighbour-joining method and the TDRAW program was used to produce the appropriate tree. The genetic distance between types is found in Table 7.2.

*this PFGE type represents 6 isolates all of which were the same ERIC type

Table 7. 3 Hybridisation patterns of the OMP26 probe among 50 Hib isolates

Group	Isolate	OMP26 fragments**	Gel
A5b	S11	29	Figure 4.3
A8b	S1	*	"
A8b	S6	13	"
A8b	S17	13	"
A8b	S18	13	"
A10b	S2	13	"
A11b	S13	13	"
A15a	S8	13	"
C1a	S14	13	"
B7	S16	13 and 77	"
A1b	S21	13 and 50	Figure 4.5
A8b	S23	13	"
A8b	S24	*	"
A8b	S25	13	"
A8b	S26	13	"
A8c	S27	*	"
A8b	S30	13	"
A8b	S28	13	"
B4	S29	13	"
C1a	S19	*	"
A1a	WA25	13 and 50	Figure 4.8
A8b	WA22	13	"
A8b	WA23	13	"
A8b	WA27	13	"
A8b	WA28	13	"
A8b	WA30	13	"
A8b	WA37	13	"
A11b	WA31	13	"
A11b	WA34	13	"
A11b	WA35	13	"
A13b	WA24	13	"
C3	WA36	13	"
A3d	V14	13 and 52	Figure 4.2
A8b	V11	13	"
A8b	V12	13	"
A8b	V13	13	"
A8b	V15	13	"
A8b	V16	13	"
A8b	V17	13	"
A8b	V20	13	"
A8b	S34	13	"
A15b	V19	13	"
B5	V18	13	"
A16	WA5	15	Figure 5.8
F2a (n=6)	WA6-10, A1	4	"

*An asterisk means that residual DNA in the loading well hybridised indicating a positive result but no fragments hybridised because the larger fragments 'washed out' in the PFGE run (this is an unexplained artifact of the procedure that occurred occasionally in some lanes on some gels; on repeat runs the presence of the large fragments were confirmed)

**numbers refer to one of the 81 marker fragments in the *Sma*I database

Presence of OMP26 gene in Hib strains Genomic DNA of all the strains (n=50) hybridised with the probe. Either one (n=46) or two (n=4) fragments per isolate hybridised. Of the 50 isolates all but 8 carried the OMP26 gene on the same size fragment. This fragment (number 13 in Table 7.3) corresponds to the 256 kb fragment of the Hib size standard, HS008, depicted in Table 2.3 and seen in each of the figures in the lane labelled 8. The fragment numbers used in Table 7.3 correspond to one of the 81 *Sma*I different sized fragments (listed numerically in the reference database) that were found among 213 Hib isolates. That is, a total of 81 different sized fragments were found among the entire collection used in the larger PFGE study. In total, in this sample of 50 isolates, seven different sized fragments hybridised with the probe.

The OMP26 hybridisation patterns were as follows:

1) one fragment, fragment 13, of all the isolates that fell into type A8b (n=22 not including S1 and S24) hybridised with the OMP26 probe,
2) the 4 isolates for which two fragments hybridised, one of which was fragment 13, belonged to 4 other RFLP types,
3) fragment 4 of all 6 isolates that fell into type F2a hybridised with the probe and,
4) two other different sized fragments hybridised in one of 2 other distinct types in which only one fragment hybridised.

7.4 Discussion

ERIC-PCR The bacterial strains selected for ERIC-PCR typing represented the genetic diversity found among Hib using PFGE. They included 6 isolates representing each of the two predominant PFGE RFLPs found in a sample of 213 Hib isolates. The findings from ERIC-PCR are in agreement with those derived from the PFGE RFLP analysis, and a correlation of 0.9 was found when a representative distance matrix derived from PFGE data was compared to that derived from the ERIC-PCR data. Calculation of Simpson's index of diversity showed that *Sma*I digestion using PFGE was more discriminating. D = 0.974 for PFGE, indicates a very high power of discrimination. The D value of ERIC-PCR at 0.890 is considered to be less desirable but is just within the range where results can be interpreted with confidence. No correction factor was used for a small sample and it is best that typing schemes be validated with larger samples ($n \geq 100$).

The major clusters estimated with PFGE were revealed with ERIC-PCR but fewer types were found among them. ERIC-PCR detected a smaller genetic distance between the two most predominant types than revealed with PFGE. The *F* value found from a pairwise comparison of F2a and A8b was 0.6250 for PFGE and 0.3448 for ERIC-PCR. This was the largest discrepancy seen in the data. Further investigation into the genetic relationship of these two isolates is warranted.

The overall high level of concordance between the two systems validates the usefulness of PFGE for typing Hib and does not contradict the use of quantitative analysis of PFGE for revealing phylogenetic relationships among the strains. The results of ERIC-PCR do not show that some strains group according to diagnosis.

OMP26 The bacterial strains selected for OMP26 hybridisation represented the predominant RFLP types, A8b (n=24) and F2a (n=6), and 20 other isolates that fell into the major clusters designated A (n=12), B (n=3) and C (n=3). Genomic DNA of all strains hybridised with the probe, the majority hybridising to the same size fragment. In every case where more than one isolate was representing a type the same size fragment hybridised. This indicates that the OMP26 gene exists in all these strains and suggests that it is found in the same chromosomal region. The results revealed the location of homologous DNA (~615 bp) on same size fragments generated by *Sma*I digestion of Hib DNA and suggest that fragments of the same size may represent the same region of the chromosome. Similar results seen with the *cap* probe and P2 probe (data not shown) when identical (but different for each probe) size fragments hybridised with the majority of isolates is further evidence of homologous DNA on same size fragments.

Concluding remark The results of the work described here do not contradict the overall genetic relationships revealed among the Hib isolates by quantitative analysis of PFGE RFLPs.

CHAPTER 8

Examination of

Isolates from the Post-vaccine Era

8.1 Introduction

One of the goals of the work presented in this thesis is to develop a database of Hib genotypes found in Australia before the introduction of conjugate vaccines and to use it to monitor Hib isolates recovered after widespread vaccination. The database has been developed by the identification of 69 genotypes (*Sma*I genotypes) among 213 Hib isolates from different population groups and different geographic areas in Australia. In this study we use the database in the examination of isolates recovered after the introduction of vaccination.

In collaboration with Heidi C. Smith-Vaughan of the Menzies School of Health Research, it was proposed to subject isolates collected in the years 1996-1998 to PFGE to study the impact of Hib vaccines on Hib genotypes endemic in a remote Aboriginal community. Isolates recovered from this same population group in the two years prior to vaccine introduction are included in the pre-vaccine database of Hib genotypes and, thus, are available for direct comparison with the post-vaccine isolates. The PFGE results of these pre-vaccine isolates are in Chapter 5 in this thesis. They comprise 19 isolates collected for a longitudinal study examining the relationship between bacterial colonisation of the nasopharynx and the onset of otitis media in Aboriginal infants (Leach *et al.*, 1994).

Prior to the introduction of the Hib conjugate polysaccharide vaccine, PRP-OMP, in mid-1993, the cumulative incidence of Hib in the infants under study was at least 42.8% by age 6 months (Smith-Vaughan *et al.*, 1998). Widespread vaccination initially resulted in a sharp decline in Hib carriage during the first six months of vaccine usage and reached zero in 1994, however, Hib carriage rates in infants subsequently rose and Hib continues to be circulated in the community (Smith-Vaughan *et al.*, 1998). This is of potential concern, particularly, since high carriage rates and a re-emergence of Hib disease has been reported among Native Alaskans (Galil *et al.*, 1999).

Temporal shifts in the subtype distribution of Hib causing invasive disease have been demonstrated (Barenkamp *et al.*, 1983; Hampton *et al.*, 1983; Leaves and Jordens, 1994); and, it is not clear what effect vaccine usage will have on the selection of Hib clones in a population that experiences continued high carriage rates. If vaccination levels fall, existing Hib clones could spread more widely. It is also theoretically possible that the uptake of capsule gene sequences by potentially invasive NCHi clones during

concurrent carriage of Hib and NCHi strains could generate new Hib strains and cause a recurrence of Hib disease (Smith-Vaughan *et al.*, 1998).

Studies on the temporal trends of subtypes of *Bordetella pertussis* have led to some intriguing observations. *Bordetella pertussis*, like Hib, is a uniquely human pathogen that was a major cause of childhood illness and infant death before the introduction of vaccination programs in industrialised countries. There has been a resurgence of pertussis (whooping cough) since the early 1980s in countries with high vaccination rates, such as, the USA, The Netherlands, Canada and Australia (Mooi *et al.*, 1998; Cordova *et al.*, 2000; Hardwick *et al.*, 2002). Pertussis vaccine was introduced in the 1940s and coverage has been continuously high in developed countries since then. The shift in epidemiology may be a consequence of a number of factors but one explanation involves changes in the circulating *B. pertussis* population leading to increased virulence or resistance to vaccine induced immunity. To investigate this, Hardwick and co-workers used PFGE to subtype recently and previously circulating *B. pertussis* isolates in the USA. They found several major trends: different types circulated and predominated at different times, the relatedness among PFGE types was consistent with a relatively homogenous population, the more frequent profiles were more highly related to each other then to less common types, and genetic diversity decreased over the study period (1935-1999) (Hardwick *et al.*, 2002).

The observed tendency of different types of *B. pertussis* circulating in pre- and post-vaccination periods was also demonstrated in the Netherlands where 83% of DNA types were limited to single periods and only one type was cultured in all successive periods (Mooi *et al.*, 1998). In particular, in the Hardwick study, PFGE types from the pre-vaccine era (before 1946) were highly divergent from the post-vaccination types. However, the role of vaccine selection in the divergence observed remains speculative because of the low number of isolates from the archival period (Hardwick *et al.*, 2002). Both of these studies provide data about the distribution of types over an extended period of time and demonstrate shifts in genetic types suggesting that the population structure of *B. pertussis* may be distinct in vaccinated and unvaccinated populations. Though the role vaccination plays in subtype selection remains equivocal, the need for continued surveillance of the genetic diversity of *B. pertussis* to explain these observations is unquestioned.

The question can be raised as to whether the introduction of vaccines will affect the types of Hib circulating in vaccinated populations in Australia, particularly those with high carriage rates. In the Aboriginal population under study a shift in β-lactamase production among Hib isolates has been observed. That is, β-lactamase production was found in 43% of pre-vaccine isolates compared to 19% of post-vaccine isolates and, if cases of possible longitudinal carriage are removed, the rate of β-lactamase production fell from 45% to 15% (personal communication, Heidi C. Smith-Vaughan). It has not been determined whether this is due to chromosomal or plasmid mediated changes in the bacteria.

Prior to the introduction of vaccines 19 Hib isolates from infants in this remote Aboriginal community were separated into 6 subtypes in two distinct lineages by PFGE typing. Here we present the results of an investigation into the possible change in the types of Hib circulating in this community since vaccination was introduced in mid-1993.

8.2 Methods

Preparation of DNA, restriction endonuclease digestion, PFGE conditions, staining and visual determination of the numbers and mobilities of fragments were performed as previously described. The gel images were captured using the Gel Doc 1000 DNA gel analysis and photodocumentation system (Bio-Rad Laboratories, Hercules, CA). The images were then pasted into Microsoft Powerpoint® for illustration and subsequent publication.

8.2.1 Bacterial isolates

Twenty Hib isolates were obtained from a collection held at the Menzies School of Health Research, Darwin, Northern Territory. They were collected in 1996-1998 as part of the longitudinal study on otitis media. They were previously identified by their requirement for X and V factor and serotyping. Serotyping was confirmed by PCR. The sample includes: 2 isolates collected from an infant and its mother on the same day (B22-23), 2 isolates from an infant and its sibling collected six months apart (B27-28), 2 isolates from the same infant collected 1 ½ years apart (B29-30), 4 isolates collected from an infant and its sibling on the same day and from this infant on two other

occasions (B31-32, B33-B34) and, 4 isolates collected from an infant and its sibling on two different occasions (B36-37 and B38-39). The other 6 isolates are from unrelated infants or siblings (B24-26, B35, B41 and B46). Table 8.1 describes the isolates.

Table 8. 1 Post-vaccine carriage isolates recovered from Aboriginal infants, siblings, and a mother from Bathurst Island

Code (PEM)	ID (MSHR)	Lab # (MSHR)	Relation	DOC	Age (days)	β-lacta-mase	PFGE type[a]	
							Group	Bathurst Is. Subset
B22	A04	4411	infant	09/02/98	644	neg	F2a	B6
B23	A04	4411	mother	09/02/98	na	neg	F2a	B6
B24	A11	4139	sibling	18/09/96	na	pos	C7a	B2
B25	A14	4269	infant	24/04/97	38	neg	C7a[b]	B25
B26	A15	4438	infant	06/05/98	1021	neg	F2a	B6
B27	A19	4354	infant	14/10/97	69	neg	B1	B1
B28	A19	4420	sibling	08/04/98	na	neg	B1[c]	B28
B29	A21	4022	infant	22/04/96	286	neg	C7a[d]	B29
B30	A21	4358	infant	14/10/97	658	neg	F2a	B6
B31	A25	4313	infant	06/08/97	53	pos	C7b	B13
B32	A25	4313	sibling	06/08/97	na	pos	C5	B8
B33	A25	4322	infant	18/08/97	65	pos	C5	B8
B34	A25	4338	infant	15/09/97	93	pos	C7a	B2
B35	A28	4374	infant	10/11/97	198	neg	F2a	B6
B36	A29	4115	infant	27/08/96	168	neg	F2a	B6
B37	A29	4115	sibling	27/08/96	na	neg	F2a	B6
B38	A29	4167	infant	15/10/96	217	neg	F2a	B6
B39	A29	4142	sibling	19/09/96	na	neg	F2a	B6
B40	A31	4272	infant	24/04/97	559	neg	B1	B1
B41	A33	4082	sibling	09/07/96	na	neg	F2a	B6

Shading is used to highlight associated isolates, i.e., B22-23, B27-28, B29-30, B31-34, and B36-39.

[a]The PFGE type is shown for the entire database grouping scheme (Group) and the subset of Bathurst Island isolates (Bathurst Is. Subset) as determined by *Sma*I RFLPs. Group and Subset designation for each isolate refer to the same RFLP. The designations are used in Figure 5.10 and Figure 6.3, the dendrograms showing the clustering of *Sma*I RFLPs in the Bathurst Island subset and the entire database, respectively.

[b]The RFLP for this isolate was not found among the pre-vaccine isolates but it is closely related to group C7a from which it differs by three fragments.

[c]The RFLP for this isolate was not found among the pre-vaccine isolates but it is closely related to group B1 from which it differs by two fragments.

[d]The RFLP for this isolate was not found among the pre-vaccine isolates but it is clonally related to group C7a from which it differs by two fragments.

8.3 Results

All the isolates produced well-resolved patterns when digested with *Sma*I and *Apa*I (see Figures 8.1–8.4). Eight *Sma*I-RFLP types were found. Five of them were indistinguishable from types found in the Hib database. The other 3 types were clonally or closely related to a type in the database. One type that was indistinguishable from the most common Aboriginal type in the Hib database, F2a, predominated in this sample accounting for one half of the isolates compared to 32% in the pre-vaccine sample. After removal of possible longitudinal carriage isolates this type accounts for 38% of the isolates in the post-vaccine sample compared to 24% in the pre-vaccine sample. Another type, C7a, accounts for 2 (10%) of the post-vaccine isolates and 7 (37%) of the pre-vaccine isolates. After removal of possible longitudinal carriage isolates there was no change in distribution among the post-vaccine isolates and 4 (26%) of the pre-vaccine isolates are type C7a. Table 8.1 lists the overall group type and the Bathurst Island subset type described previously in this thesis for each of the isolates.

The distribution of isolates among the types in both the post-vaccine and pre-vaccine samples was:

Northern Territory Bathurst Island post-vaccine isolates (*n=20*) compared to pre-vaccine isolates (*n=19*) by *Sma*I-RFLP type

(Pre-vaccine) *Sma*I-RFLP type Bathurst Is. Subset type / Group type	Number of isolates per type		Number of isolates per type after removal of possible longitudinal carriage isolates	
	Post-vaccine number (%)	Pre-vaccine number (%)	Post-vaccine number (%)	Pre-vaccine number (%)
B1 / B1	2 (10)	1		
B2 / C7a	2 (10)	7 (37)	2 (10)	4 (25)
B5 / F2c	0	1		
B6 / F2a	10 (50)	6 (32)	6 (38)	4 (24)
B8 / C5	2 (10)	3 (16)		
B13 / C7b	1	1		
B25 / new[a]	1	0		
B28 / new[b]	1	0		
B29 / new[a]	1	0		

[a]The RFLP of this isolate was not found in the pre-vaccine sample but is closely related to pre-vaccine type C7a.
[b]The RFLP of this isolate was not found in the pre-vaccine sample but is clonally related to pre-vaccine type B1.

The use of the restriction enzyme *Apa*I revealed seven *Apa*I-RFLP types in 3 clonal groups. The 10 isolates indistinguishable from *Sma*I-RFLP F2a were broken in to 3 *Apa*I-RFLP types that differed from each other by 1 fragment.

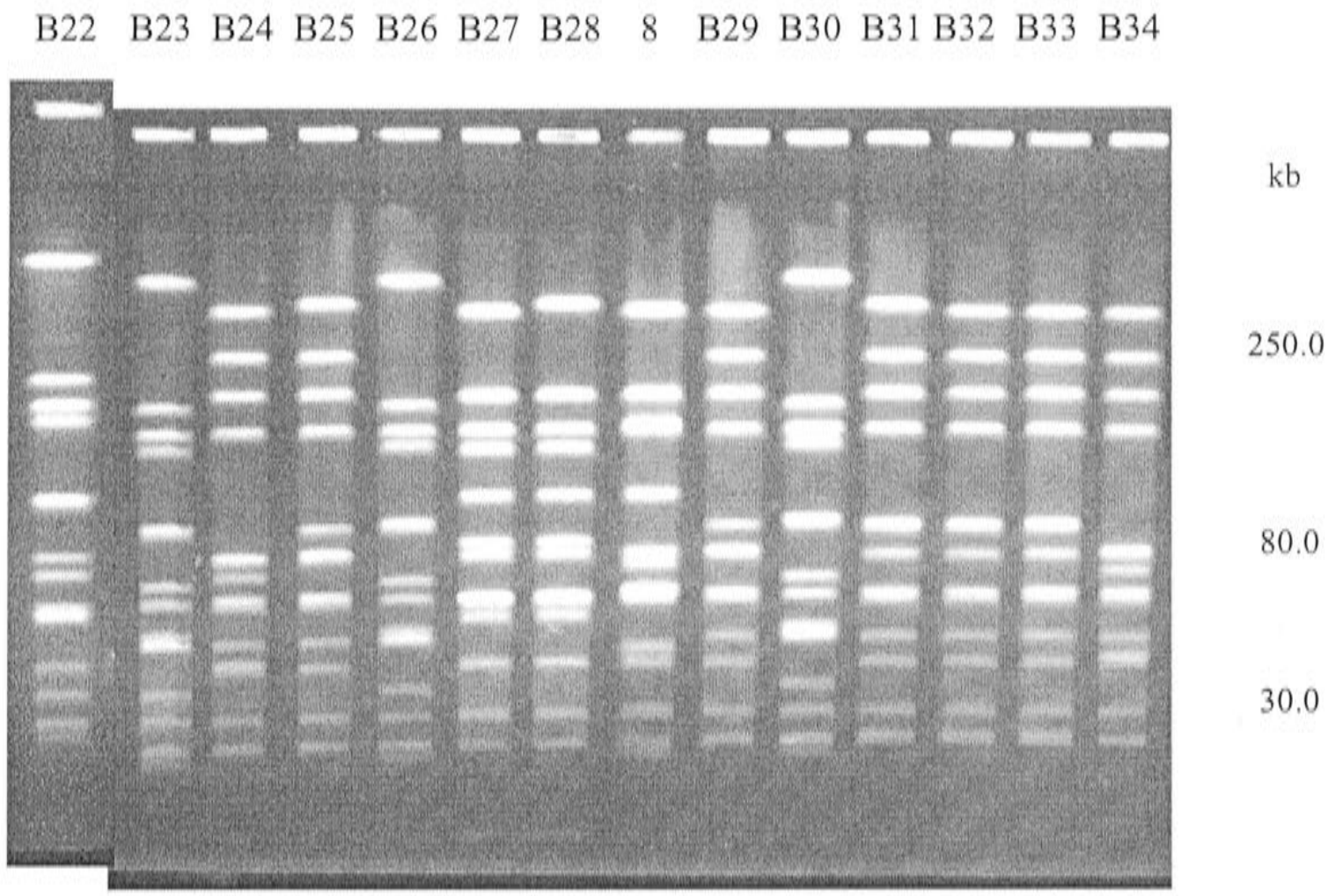


Figure 8. 1 *Sma*I restriction fragment patterns of post-vaccine isolates

*Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 13 Aboriginal Hib isolates from Bathurst Island that were collected after the introduction of widespread vaccination. The lane labelled 8 contains fragment size standards. (Note: Lane B22 has been cut and pasted into this picture from an image of another gel run.)

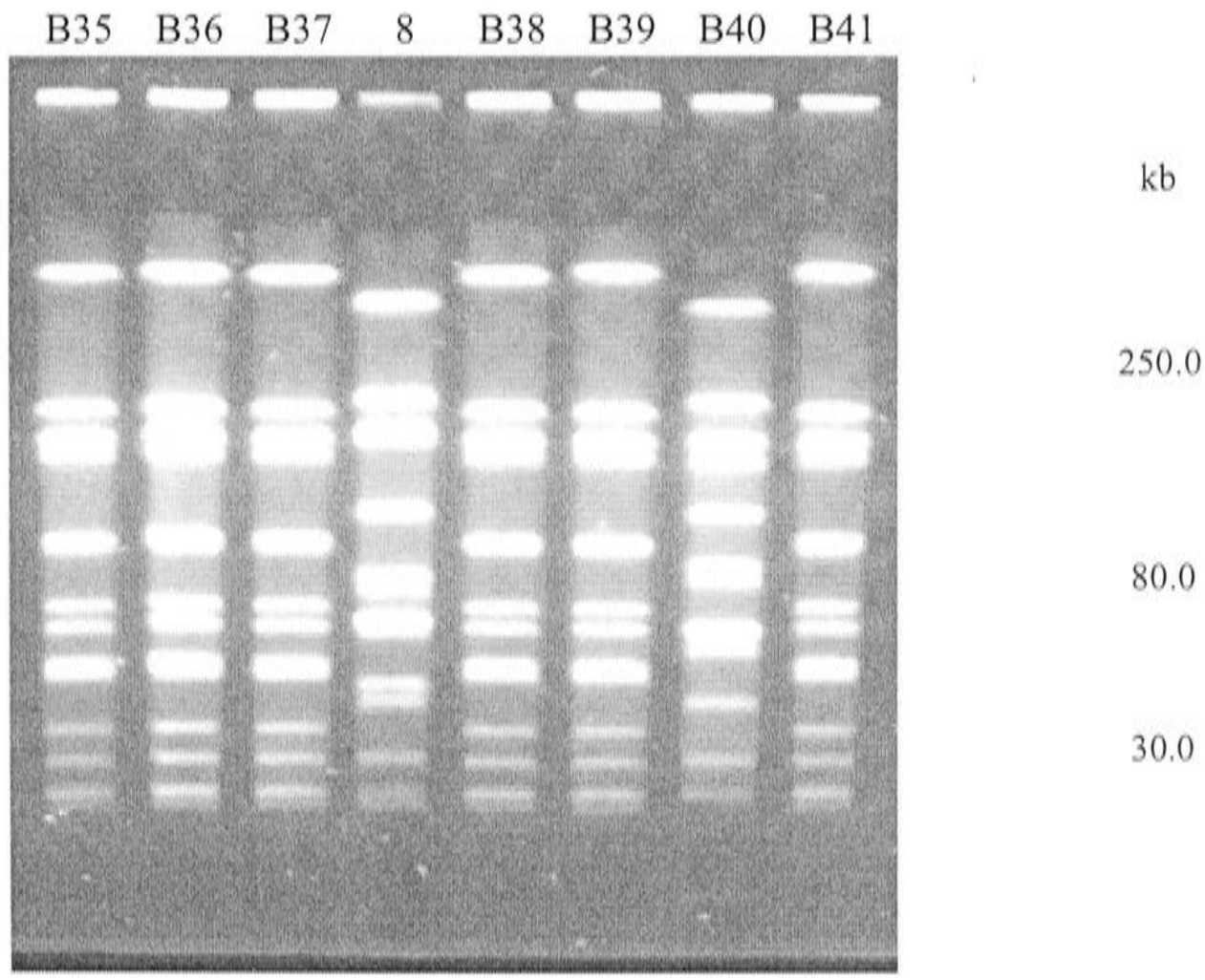


Figure 8. 2 *Sma*I restriction fragment patterns of post-vaccine isolates

*Sma*I restriction fragments obtained by PFGE of the genomic DNAs of 7 Aboriginal Hib isolates from Bathurst Island that were collected after the introduction of widespread vaccination. The lane labelled 8 contains fragment size standards.

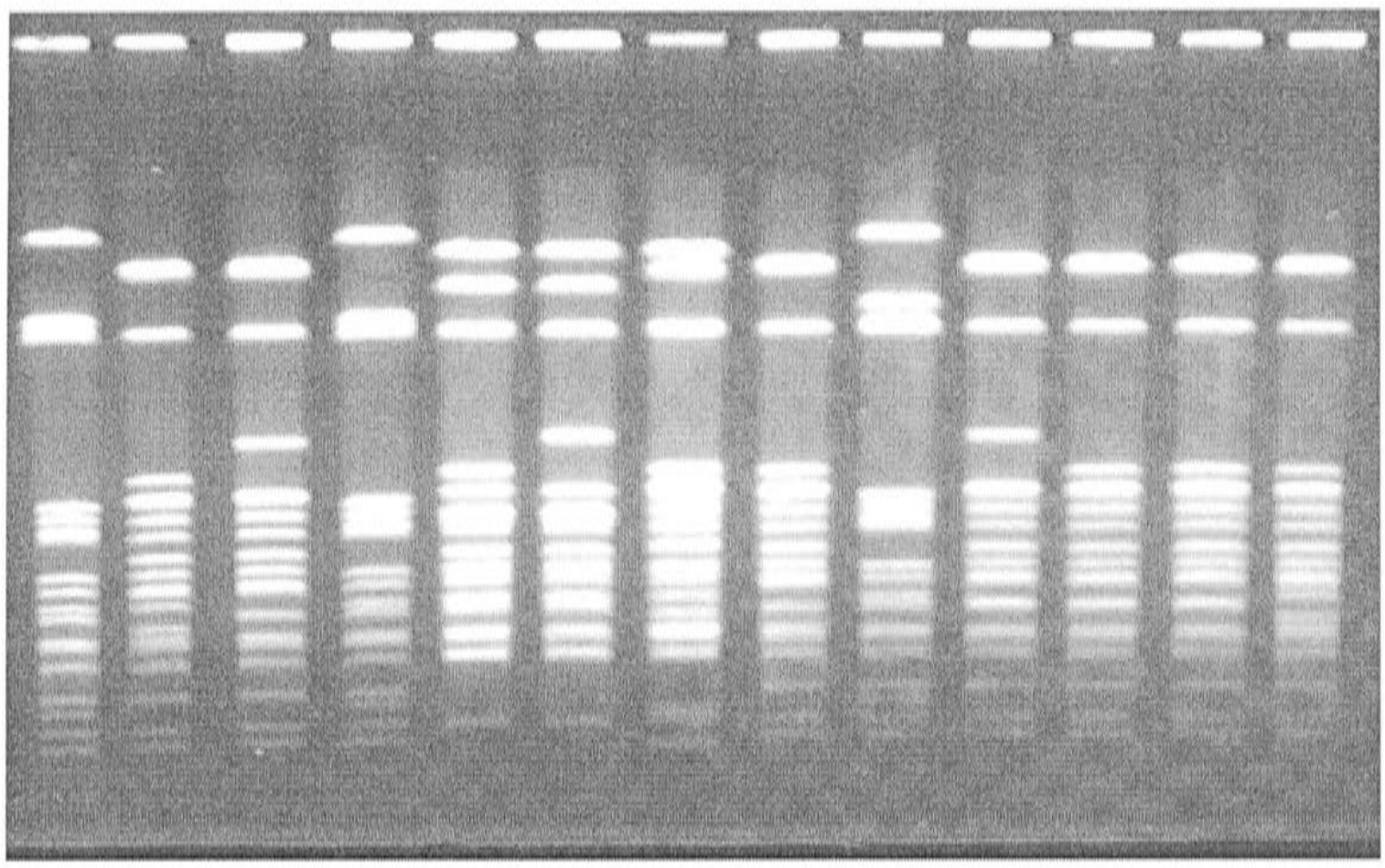


Figure 8. 3 *Apa*I restriction fragment patterns of post-vaccine isolates

*Apa*I restriction fragments obtained by PFGE of the genomic DNAs of 12 Aboriginal Hib isolates from Bathurst Island that were collected after the introduction of widespread vaccination. The lane labelled 8 contains fragment size standards.

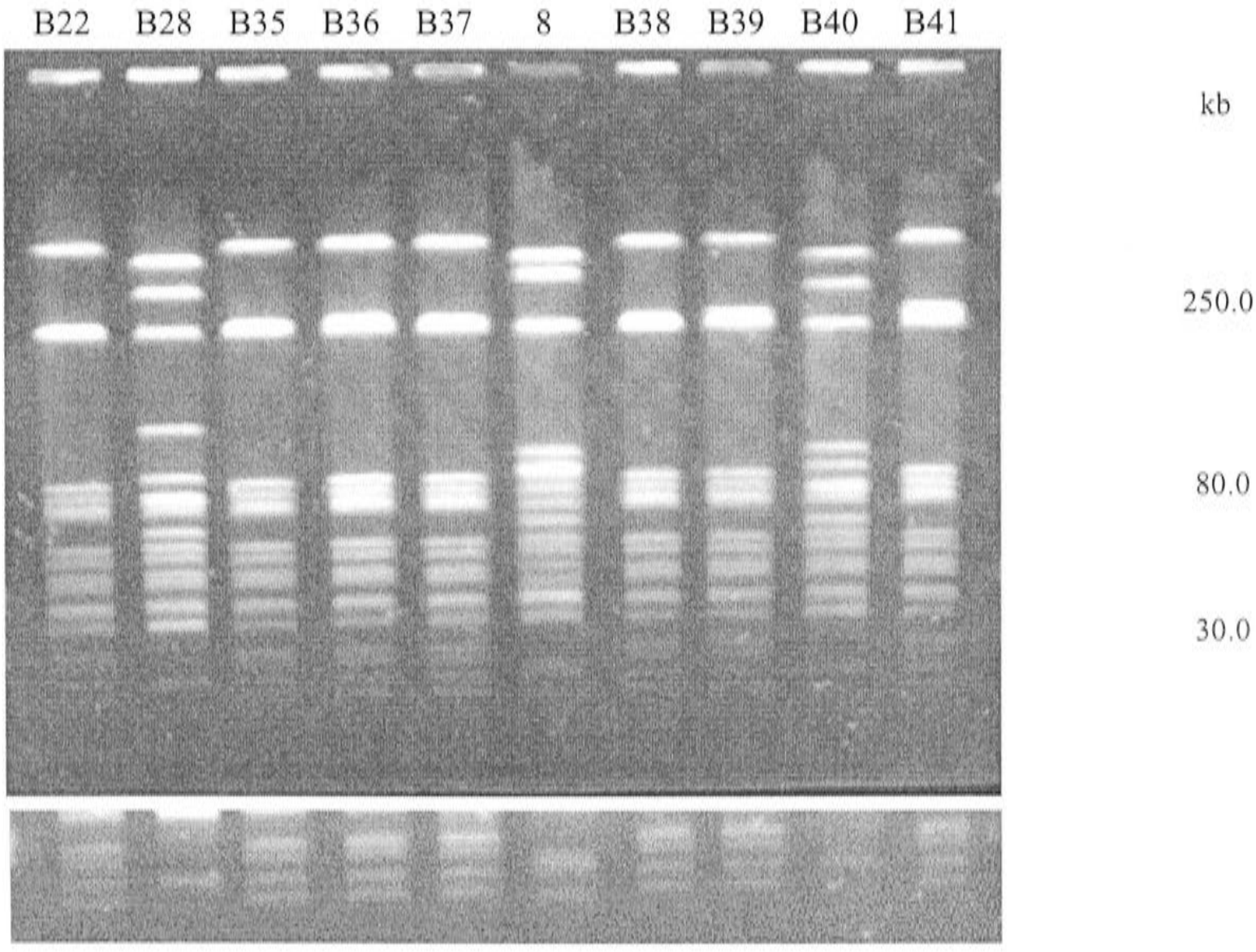


Figure 8. 4 *Apa*I restriction fragment patterns of post-vaccine isolates

*Apa*I restriction fragments obtained by PFGE of the genomic DNAs of 9 Aboriginal Hib isolates from Bathurst Island that were collected after the introduction of widespread vaccination. The lane labelled 8 contains fragment size standards. (Note: The smallest fragments are shown to better advantage in the section appended to the bottom of this depiction.)

8.4 Discussion

There has been a resurgence of Hib carriage in a remote Aboriginal community since widespread vaccination introduced in the early 1990s initially caused a sharp decline in its carriage. Because this resurgence may represent changes in the carriage isolates, we determined the PFGE profiles of 20 isolates circulating in the population in this community after the introduction of Hib vaccines and compared them to PFGE profiles of Hib isolates recovered in the same community before the introduction of vaccines. We looked for any trends with potential epidemiologic significance.

We found that the same PFGE types were circulating during the pre-vaccine era and several years after the introduction of vaccines. Of 8 *Sma*I-RFLP types found among the post-vaccine isolates 5 are indistinguishable from types found in the pre-vaccine isolates from the Bathurst Island community under study. The same type (indistinguishable from Group F2a) predominates in both the pre- and post-vaccine samples and there is no evidence of displacement of types in the post-vaccine sample. Three of the post-vaccine types are not found in the pre-vaccine sample from the community nor from the Australia wide sample but they are clonally (differed by 1 fragment) or closely (differed by 2-3 fragments) related to two of the other 5 post-vaccine types.

The divergence among these 3 types in the post-vaccine sample compared to the pre-vaccine sample may suggest a shift in population structure but because of the relatively low number of isolates in each sample it cannot be ruled out that less frequent types were not recovered in the pre-vaccine sample.

PFGE is particularly suited to detecting micro-evolutionary changes and the results found in this small sample suggest that the Hib genomic types circulating several years after the introduction of vaccines are similar to those circulating before immunisation. Because of the pre-vaccine sample size the data does not suggest that new variants of Hib are present after the resurgence of carriage in this population but the detection of 3 new closely related types warrants investigation. Examination of a larger sample of isolates and further characterisation of isolates using other molecular methods that reveal the DNA sequences within fragments may reveal the significance of these variants.

Hib carriage is a dynamic process. It was shown in another sample of isolates from this same population that ERIC-PCR typing split the Hib isolates into two groups with the apparent replacement of one variant by the other during several episodes of Hib carriage (Smith-Vaughan *et al.*, 1998). It is not clear whether the two variants of Hib were transmitted, and thus, carried together, taking turns as the dominant strain, or whether there were separate episodes of colonisation in infants who do not develop a strong immune response to the Hib polysaccharide capsule following natural exposure.

PFGE also revealed the dynamics of Hib carriage. In this study:

1) the same PFGE type was recovered from a mother (B22) and her infant (B23) on the same day,
2) the same PFGE type was recovered from an infant (B36, B37) and its sibling (B38, B30) on two occasions within a 2 month period,
3) a type recovered from its sibling (B28) was closely related to a isolate recovered from an infant (B27) six months previously,
4) three closely related types were recovered from an infant (B31, B33, B34) and its sibling (B32) over a period of 2 months and each of the types was carried by the infant at different times and,
5) two genetically distinct types collected a year and a half apart were recovered from the same infant (B29, B30).

Thus, our results show that both concurrent carriage of the same or closely related Hib types among relatives occurs over a period of at least two months and that an individual may harbor closely related or distinctly different types at different points in time. It is not clear whether the subsequent presence of a closely related type represents colonisation from a newly introduced strain or is a strain carried concurrently and not previously recovered due to sampling methods or is a recent variant of the previously identified isolate to which it is related. The recovery of distinctly different types from an infant and its sibling implies that the Hib isolates were from different origins and that older siblings were not the source of the infant Hib isolate.

Although the *Sma*I restriction profiles of B27 and B28 that were recovered from an infant and its sibling were almost identical they differed by two fragments. When the genomic DNAs of B27 and B28 were digested with *Apa*I a two fragment difference was

also observed. Since the difference observed with *Sma*I could also be observed with *Apa*I, this meant that the band differences are not simply due to a point mutation that modified the *Sma*I restriction site. This may indicate that the strains isolated from these siblings were of different origins.

Concluding remark This work demonstrates the usefulness of the PFGE database developed during the course of the larger study for the evaluation of Hib isolates recovered after the introduction of vaccines.

CHAPTER 9

Concluding Remarks

Before the introduction of conjugate vaccines in 1992, *Haemophilus influenzae* type b (Hib) disease was a major cause of morbidity and mortality in Australian children particularly among Aboriginal populations in whom the incidence of Hib disease was among the highest in the world. Overall 600-700 cases of invasive Hib disease occurred each year in Australia. While the incidence of Hib disease has dropped dramatically since vaccination policies have been in place, invasive disease due to Hib still occurs. Between July 1993 and June 1996, 412 cases, including 18 deaths, were reported to the Hib Case Surveillance Scheme. Thirty-four cases met the Australian case definition of a vaccine failure. However, a further 24 cases would meet the United Kingdom vaccine failure definition that includes cases occurring after just two rather than three doses of vaccine when the first dose is given before the age of 7 months (Herceg, 1997). That Hib continues to circulate in an Aboriginal island community 6 years after the introduction of the Hib conjugate polysaccharide vaccine is a cause for concern, particularly if vaccination rates fall (Smith-Vaughan *et al.*, 1998). Thus, despite the introduction of effective vaccines in 1992 there is a public health need to monitor Hib and other invasive *H. influenzae* populations in Australia.

The overall purpose of this project was to characterise the population structure of Hib in Australia using PFGE, determine whether RFLP types correlated with the epidemiology of disease, and assess the usefulness of PFGE as the basis of a database of genetic types to monitor Hib isolates in the post-vaccine era. This has been accomplished.

The ultimate collection of isolates made available for this study is the result of professional networking rather than any well laid out plan that would ensure an ideal selection. Many of the isolates had been collected for a nationwide multi-city study on antimicrobial resistance but others came from contacts who collected Hib for a variety of purposes. Clinical and non-clinical isolates were obtained. The history that was available for each isolate varied considerably and though all isolates could be separated into carrier or disease associated types, in some cases, for a few invasive isolates a specific diagnosis was not able to be obtained. Missing data were always thoroughly investigated with the source person from whom I obtained isolates. Sometimes it was just not available. It should be noted that the isolates were collected over a period of a few years as I became acquainted with the microbiology community in Australia. The point is, I did not have immediate access to isolates recovered in most of the varied geographic regions that were ultimately included in this study. I was responsible for

locating and obtaining all the isolates except the Canberra and Sydney strains that are stored at the Canberra Hospital. The quality of the sample did not suffer too much from this ad hoc collecting. A particularly sorry point, though, is that I was not able to obtain Aboriginal isolates from urban areas and thus could not directly address a very important question, namely, are Hib clones like those described elsewhere associated with specific geographic areas?

The work reported here established the usefullness of PFGE for examining the genetic diversity of Hib in Australia. The practical applicability of this for surveillance was shown when 2 isolates of Hib recovered in 2 fatal cases in a nursing home outbreak in the Sydney region (Heath *et al.*, 1997) were analysed, and both were shown to have patterns identical to that of *Sma*I type A8b. The database was also helpful in confirming the identity of a nontypeable *H. influenzae* from the cerebrospinal fluid and blood of a child with a past history of Hib meningitis and subsequent immunisation (Moor and Kelly, 1995). Medical and laboratory evaluation failed to find an immunological or anatomical predisposition for this child to suffer recurrent invasive infection. Southern blot analysis confirmed the lack of an intact capsulation (*cap*) locus within the chromosome showing that this isolate was not a capsule deficient type b. RFLP analysis using PFGE further confirmed the strain was not closely related to known Hib isolates previously typed by this method. The unusual presentation was also notable for the recovery of a serologically typed Hib isolate from a nasal swab taken at the time of the NTHi meningitis. Unfortunately, this isolate was not available for analysis.

PFGE and the Australia wide database is now being applied to the examination of Hib carrier strains found in the Bathurst Island Aboriginal community after the introduction of vaccines. The emergence of multiresistant Hib in Papua New Guinea (PNG) is of some concern (Michael *et al.*, 1997). Five new types were found among 7 isolates sent to me from patients at Goroka Base Hospital PNG to evaluate for genetic diversity. The data revealed that the isolates were not the result of the clonal spread of a single type.

The widespread introduction of vaccines against *Haemophilus influenzae* type b (Hib) increases the importance of the precise identification of isolates infecting both immunised and non-immunised patients. There is an increasing need to more precisely characterise the genetic characteristics of NTHi as they become more prominent in the infections caused by *H. influenzae*. Preliminary work has begun on a PFGE RFLP

database of NTHi modeled on this system (Webb-Wagg, 1997) and a project using other molecular typing methods in conjunction with PFGE to characterise NTHi in Australia is under development.

Since the commencement of this project others have reported using CHEF PFGE to study smaller collections of Hib isolates (Tarasi *et al.*, 1998; Mitsuda *et al.*, 1999; Saito *et al.*, 1999) and their results have also confirmed the effectiveness of PFGE for studying the genetic diversity of Hib. As noted, what was an almost experimental technique when I began this study is now the gold standard of typing.

There is real concern about making phylogenetic inferences with RFLP data and it must be emphasised that the values of *F* obtained should not be seen as precise estimates of genetic distance but summary values indicative of overall similarities and differences between isolates and a means to illustrate relationships within a large body of comparative data. The results of the *cap* locus, P2 and OMP26 hybridisation studies showed that similar size fragments carried the same genetic information. Thus, the results of these tests did not contradict the inferences made by RFLP analysis and revealed homologous DNA was carried on same sized fragments. By definition Hib isolates carry the *cap* locus but the analyses also showed that the P2 and OMP26 genes are highly conserved among the different isolates.

The results described in this thesis are based largely on the RFLPs of a single enzyme that was shown to correlate very well with a second rare cutting endonuclease. The use of more than one enzyme would allow sampling of other parts of the genome and would theoretically result in a more powerful analysis. Other enzymes may detect molecular differences that demonstrate clonal disease specificity that was not found with *Sma*I or *Apa*I. This should continue to be investigated. Not only, to more carefully examine the diversity of Hib, but to better understand the RFLP data generated by PFGE.

PFGE is currently considered the most reliable and practical tool for molecular epidemiologic analyses of bacterial infections and it is particularly well suited to detecting microevolution. But, we agree that additional analysis using other genotyping techniques should be used in conjunction with RFLP data. Rapid amplified polymorphic DNA (RAPD) analysis was shown to improve the subtyping of 30 Hib strains characterised by PFGE in a recent Japanese study (Mitsuda *et al.*, 1999). Evaluation of

other genotyping systems and combinations of them with PFGE should be explored. Polymerase chain reaction-based typing systems, particularly RAPD RFLP analysis are now widely used for bacterial typing and have been compared to PFGE (Chachaty *et al.*, 1994; Farber and Addison, 1994; Sandery *et al.*, 1994; Renders *et al.*, 1996). The results of ERIC-PCR used in this study were highly correlated with the PFGE results and did not contradict the phylogenetic inferences made with PFGE. Further evaluation of the purported relationships should be explored.

Long PCR-ribotyping was found to be as informative for typing nontypeable *H. influenzae* as conventional ribotyping and subsequently has been used to type Hib isolates (Smith-Vaughan *et al.*, 1995; Smith-Vaughan *et al.*, 1998). Although sampling a highly conserved region of the genome, PCR ribotyping may also provide further informative genetic characterisation of Hib among a large diverse sample of isolates. The availability of molecular genetic technologies like DNA sequencing, and physical and genetic map construction would also permit a clearer insight into the local population structures of Hib.

Prior to the widespread use of molecular techniques to characterise bacteria, phenotypic methods, such as, OMP typing and multilocus enzyme electrophoresis (MLEE), were used extensively in the study of Hib. Examination of the sample studied here with these two typing systems (a formidable task in the case of MLEE) would allow an invaluable comparison with isolates previously typed by these methods, as well as, increase the resolution of the overall analysis. MLST represents the newest generation of molecular technology and in companion with PFGE may provide the most readily available window on population structures.

The visual analysis of the gel patterns becomes laborious and immensely time consuming when more than a few isolates are compared. However, the development of computerised gel scanning and pattern analysis methods have made it possible to create reference databases in a fraction of the time (Gerner-Smidt *et al.*, 1998).

The analyses described in this thesis have given us a better understanding of the population structure of *Haemophilus influenzae* type b in Australia. OMP subtyping revealed that a small number of OMP subtypes caused most Hib disease but OMP subtyping is not very discriminatory and did not reveal the diversity seen with PFGE.

The data confirmed that the population structure of Hib in Australia is clonal and revealed that two different types, designated A8b and F2a, predominated among Aboriginal versus non-Aboriginal children. A deep divergence in the genetic relatedness of Aboriginal and non-Aboriginal isolates was revealed. Because of the increased incidence of Hib disease in Aboriginal populations it is tempting to speculate on whether this large genetic difference is associated with increased virulence among the Aboriginal Hib types. According to what I call Musser *et al.*'s (1990) 'geographic isolation' hypothesis these isolates may represent an older pattern of differentiation that has not been obscured by human migration patterns of the last few hundred years. The data was in agreement with the prediction based on this hypothesis that within a native population a few distinct clones would be found.

Important information about the dynamics of Hib strain diversity in the pre- and post-vaccine periods was revealed. Pre-vaccine PFGE types were found to be circulating several years after the introduction of vaccines in a remote Aboriginal community. Of 8 *Sma*I-RFLP types found among 20 post-vaccine isolates 5 are indistinguishable from types found among the pre-vaccine isolates. The same type (F2a) predominates in both the pre- and post-vaccine samples. The three post-vaccine types not found in the pre-vaccine sample are closely related to two of the pre-vaccine types. This divergence may suggest a shift in population structure but because of the relatively low number of isolates studied it cannot be ruled out that less frequent types were not recovered in the pre-vaccine sample.

It is vital to be able to predict the likely effects of public health interventions on the biology of the pathogens against which they are directed. Understanding the structure and dynamics of microbial populations will permit the design of effective public health policies that minimise the possibilities for microbial evolution to neutralise them. The genotypic database compiled in this study is unique in its scope and can be used in a broader national surveillance system to support this aim. No other study has examined the genotypic diversity of such a wide-ranging collection of Australian isolates and provided information on the dynamics of Hib diversity in the pre- and post-vaccine era.

The results of this work can be applied to the continued study of the epidemiology of Hib and to elucidating the genetic virulence determinants of its ability to cause disease.

APPENDICES

Appendix A.1: Hib Isolates Recovered from the Canberra Region

Code	ID #	Location	Hospital	Diagnosis	Specimen	D O C	D O B	Age(y.m.)	Sex	Race	Serotype	Biotype	B-Lactamase	pU082 "b"
M1	HS0358	Canberra	WVH	meningitis	bld/csf	2-Aug-88	not known	3	m	N	b	I	positive	positive
M2	HS0359	Canberra	WVH	meningitis	bld/csf	1-Sep-88	not known	2.2	f	N	b	I	positive	positive
M3	HS0360	Canberra	WVH	meningitis	bld/csf	11-Jun-89	not known	27	m	N	b	I	positive	positive
M4	HS0363	Canberra	WVH	meningitis	blood	20-Dec-88	not known	45	f	N	b	I	positive	positive
M5	HS0392	Canberra	WVH	meningitis	csf	18-Sep-80	not known	1.4	f	N	b	I	negative	positive
M6	HS0485	Canberra	WVH	meningitis	bld/csf	25-Feb-89	not known	1.5	m	N	b	I	positive	positive
M7	HS0486	Canberra	WVH	meningitis	bld/csf	5-Jun-89	not known	1.5	m	N	b	I	positive	positive
M8	HS0488	Canberra	WVH	meningitis	bld/csf	17-Jun-89	not known	1.9	f	N	b	II	negative	positive
M9	HS0494	Canberra	WVH	meningitis	csf	11-Apr-89	not known	1.2	m	N	b	II	negative	positive
M10	HS0495	Canberra	WVH	meningitis	bld/csf	15-Nov-89	not known	0.3	m	N	b	I	positive	positive
M11	HS0497	Canberra	WVH	meningitis	bld/csf	8-Jan-90	not known	2.1	f	N	b	I	negative	positive
M12	HS0501	Canberra	WVH	meningitis	bld/csf	21-Feb-90	not known	1.1	m	N	b	I	negative	positive
M13	HS0514	Goulbourn	WVH	meningitis	bld/csf	10-May-90	not known	1.7	f	N	b	I	negative	positive
M14	HS0515	Queanbeyan	WVH	meningitis	csf	25-May-90	not known	2.7	f	N	b	I	positive	positive
M15	HS0550	Canberra	WVH	meningitis	bld/csf	7-Jun-90	not known	0.6	m	N	b	II	negative	positive
M16	HS0551	Canberra	WVH	meningitis	bld/csf	11-Jun-90	not known	0.5	m	N	b	I	negative	positive
M17	HS0553	Canberra	WVH	meningitis	bld/csf	25-Jun-90	not known	2.9	f	N	b	I	negative	positive
M18	HS0593	Canberra	WVH	meningitis	bld/csf	17-Jul-90	not known	0.6	m	N	b	I	negative	positive
M19	HS0594	Canberra	WVH	meningitis	bld/csf	30-Jul-90	not known	1.5	f	N	b	I	positive	positive
M20	HS0598	Canberra	WVH	meningitis	bld/csf	21-Sep-90	not known	2.6	f	N	b	I	positive	positive
E1	HS0354	Canberra	WVH	epiglottitis	bld	23-Jun-88	not known	2.5	m	N	b	I	negative	positive
E2	HS0357	Canberra	WVH	epiglottitis	bld	2-Aug-88	not known	2.9	f	N	b	I	negative	positive
E3	HS0362	Canberra	WVH	epiglottitis	bld	18-Dec-88	not known	4.6	f	N	b	I	positive	positive
E4	HS0366	Canberra	WVH	epiglottitis	bld	14-Mar-89	not known	4.2	m	N	b	I	positive	positive
E5	HS0367	Canberra	WVH	epiglottitis	bld	18-Mar-89	not known	2.8	f	N	b	I	negative	positive
E6	HS0369	Canberra	WVH	epiglottitis	bld	16-Mar-89	not known	2.7	m	N	b	I	positive	positive
E7	HS0489	Canberra	WVH	epiglottitis	bld	12-Jul-89	not known	2.4	m	N	b	II	negative	positive
E8isS37	HS0490	Campbelltown	WVH	epiglottitis	bld	14-Jul-89	not known	2.6	m	N	b	I	negative	positive
E9	HS0504	Canberra	WVH	epiglottitis	bld	2-Mar-90	not known	0.4	m	N	b	I	negative	positive
E10	HS0552	Canberra	WVH	epiglottitis	bld	13-Jun-90	not known	3	m	N	b	I	negative	positive
E11	HS0554	Canberra	WVH	epiglottitis	bld	23-Jun-90	not known	2.2	f	N	b	I	negative	positive
E12	HS0595	Canberra	WVH	epiglottitis	bld	17-Aug-90	not known	3.5	m	N	b	I	negative	positive
E13	HS0596	Canberra	WVH	epiglottitis	bld	26-Aug-90	not known	3.8	m	N	b	I	positive	positive
E14	HS0599	Canberra	WVH	epiglottitis	nk	21-Sep-90	16-Feb-60	30	?	N	b	I	negative	positive
E15	HS0601	Canberra	WVH	epiglottitis	nk	14-Dec-90	24-Dec-87	2.11	m	N	b	I	positive	positive

Appendix A.2 (page 1 of 2): Hib Isolates Recovered from the Sydney Region and Townsville

Code	ID #	Location	Hospital	Diagnosis	Specimen	D O C	D O B	Age (y.m)	Sex	Race	Serotype	Biotype	B-Lactamase	pU082 "b"
S1	HS0083	Sydney	Syd Children's	cellulitis	bld	26-Sep-88	not known	0.7	f	N	b	I	negative	positive
S2	HS0084	Sydney	Syd Children's	meningitis	csf	5-Dec-88	not known	5	m	N	b	I	negative	positive
S3	HS0513	Sydney	Blacktown	invasive	bld	23-Oct-89	not known	4.3	m	N	b	I	negative	positive
S4	HS0087	Sydney	Syd Children's	cellulitis	bld	17-Nov-88	not known	1	m	N	b	I	positive	positive
S5	HS0096	Sydney	Nepean	meningitis	csf	not known	not known	1.7	f	N	b	I	negative	positive
S6	HS0116	Sydney	St. George	meningitis	bld	15-Mar-89	not known	3	m	N	b	I	negative	positive
S7	HS0117	Sydney	Nepean	meningitis	csf	8-Apr-89	not known	16	m	N	b	I	negative	positive
S8	HS0118	Sydney	Nepean	meningitis	csf	13-Apr-89	not known	5	m	N	b	I	negative	positive
S9	HS0165	Sydney	St. George	meningitis	csf	4-May-89	not known	1	m	N	b	I	negative	positive
S10	HS0207	Sydney	Blue Mountains	invasive	bld	11-Jul-89	not known	2.5	f	N	b	I	positive	positive
S11	HS0208	Sydney	St. George	meningitis	csf	13-Jul-89	not known	1.5	?	N	b	I	positive	positive
S12	HS0209	Sydney	Bankstown	meningitis	csf	17-Jul-89	not known	1.7	?	N	b	I	positive	positive
S13	HS0210	Sydney	St. George	invasive	bld	3-Ju;-89	not known	5.9	f	N	b	I	negative	positive
S14	HS0211	Sydney	Sutherland	meningitis	csf	15-Aug-89	not known	1.8	f	N	b	I	negative	positive
S15	HS0212	Sydney	Hawkesbury	invasive	bld	15-Aug-89	not known	4.1	m	N	b	I	positive	positive
S16	HS0213	Sydney	Hawkesbury	meningitis	csf	16-Aug-89	not known	1.4	f	N	b	I	positive	positive
S17	HS0215	Sydney	Westmead	meningitis	csf	21-Aug-89	not known	0.2	m	N	b	I	negative	positive
S18	HS0235	Sydney	Westmead	meningitis	csf	18-Aug-89	not known	0.5	m	N	b	I	negative	positive
S19	HS0236	Sydney	Mona Vale	invasive	bld	21-Aug-89	not known	4.8	m	N	b	I	negative	positive
S20	HS0237	Sydney	Blacktown	meningitis	csf	23-Aug-89	not known	0.7	m	N	b	I	negative	positive
S21	HS0238	Sydney	Blacktown	invasive	bld	15-Aug-89	not known	2	m	N	b	II	positive	positive
S22	HS0239	Sydney	Mt. Druitt	meningitis	csf	30-Aug-89	not known	0.6	f	N	b	I	positive	positive
S23	HS0240	Sydney	St. George	invasive	bld	1-Sep-89	not known	0.8	m	N	b	I	negative	positive
S24	HS0270	Sydney	Westmead	meningitis	csf	12-Apr-89	not known	0.2	?	N	b	I	negative	positive
S25	HS0271	Sydney	Gosford	meningitis	csf/bld	14-Apr-89	not known	0.5	f	N	b	I	negative	positive
S26	HS0272	Sydney	Bankstown	invasive	bld	8-Apr-89	not known	4.7	f	N	b	I	negative	positive
S27	HS0273	Sydney	Westmead	invasive	bld	29-Apr-89	not known	not known	f	N	b	II	negative	positive
S28	HS0274	Sydney	Gosford	meningitis	csf	16-May-89	not known	1	f	N	b	I	negative	positive
S29	HS0275	Sydney	Nepean	meningitis	csf	29-May-89	not known	1.9	f	N	b	I	negative	positive
S30	HS0276	Sydney	Nepean	meningitis	csf	6-Jun-89	not known	1	f	N	b	I	negative	positive

Appendix A.2 (page 2 of 2): Hib Isolates Recovered from the Sydney Region and Townsville

Code	ID #	Location	Hospital	Diagnosis	Specimen	D O C	D O B	Age(y.m.)	Sex	Race	Serotype	Biotype	B-Lactamase	pU082 "b"
S31	HS0277	Sydney	Westmead	invasive	bld	10-Jun-89	not known	2.6	f	N	b	I	negative	positive
S32	HS0278	Sydney	Bankstown	invasive	bld	9-Jun-89	not known	1.8	f	N	b	I	negative	positive
S33	HS0391	Sydney	Nepean	invasive	bld	5-Sep-89	not known	0.8	f	N	b	I	negative	positive
S34	HS0505	Sydney	Westmead	invasive	bld	14-Oct-89	not known	4.4	m	N	b	I	negative	positive
S35	HS0510	Sydney	Westmead	meningitis	csf	23-Sep-89	not known	0.5	f	N	b	I	negative	positive
S36	HS0512	Sydney	Hornsby	meningitis	csf/bld	15-Oct-89	not known	3.8	f	N	b	I	negative	positive
S37	HS0490	Campbelltown	WVH	epiglottitis	bld	14-Jul-89	not known	2.6	m	N	b	I	negative	positive
T1		Townsville	T'ville Gen Ho	meningitis	csf/bld	after July 93	not known	nk	nk	N	b	nk	nk	positive
T2		Townsville	T'ville Gen Ho	meningitis	csf/bld	after July 93	not known	nk	nk	N	b	nk	nk	positive

Appendix A.3: Hib Isolates Recovered from the Melbourne Region

Code	ID #	Location	Hospital	Diagnosis	Specimen	D O C	D O B	Age (y.m)	Sex	Race	Serotype	pU082 "b"
V1	1	Melbourne	R C H	meningitis	not known	1987-1990	not known	nk	nk	N	b	positive
V2	7	Melbourne	R C H	meningitis	not known	1987-1990	not known	nk	nk	N	b	positive
V3	8	Melbourne	R C H	meningitis	not known	1987-1990	not known	nk	nk	N	b	positive
V4	15	Melbourne	R C H	meningitis	not known	1987-1990	not known	nk	nk	N	b	positive
V5	35	Melbourne	R C H	meningitis	not known	1987-1990	not known	nk	nk	N	b	positive
V6	46	Melbourne	R C H	meningitis	not known	1987-1990	not known	nk	nk	N	b	positive
V7	41	Melbourne	R C H	meningitis	not known	1987-1990	not known	nk	nk	N	b	positive
V8	68	Melbourne	R C H	meningitis	not known	1987-1990	not known	nk	nk	N	b	positive
V9	95	Melbourne	R C H	meningitis	not known	1987-1990	not known	nk	nk	N	b	positive
V10	121	Melbourne	R C H	meningitis	not known	1987-1990	not known	nk	nk	N	b	positive
V11	4	Melbourne	R C H	epiglottitis	not known	1987-1990	not known	nk	nk	N	b	positive
V12	24	Melbourne	R C H	epiglottitis	not known	1987-1990	not known	nk	nk	N	b	positive
V13	12	Melbourne	R C H	epiglottitis	not known	1987-1990	not known	nk	nk	N	b	positive
V14	16	Melbourne	R C H	epiglottitis	not known	1987-1990	not known	nk	nk	N	b	positive
V15	65	Melbourne	R C H	epiglottitis	not known	1987-1990	not known	nk	nk	N	b	positive
V16	20	Melbourne	R C H	epiglottitis	not known	1987-1990	not known	nk	nk	N	b	positive
V17	29	Melbourne	R C H	epiglottitis	not known	1987-1990	not known	nk	nk	N	b	positive
V18	42	Melbourne	R C H	epiglottitis	not known	1987-1990	not known	nk	nk	N	b	positive
V19	53	Melbourne	R C H	epiglottitis	not known	1987-1990	not known	nk	nk	N	b	positive
V20	70	Melbourne	R C H	epiglottitis	not known	1987-1990	not known	nk	nk	N	b	positive

Appendix A.4 (page 1 of 1): Hib Isolates Recovered from Metropolitan Perth and Rural Western Australia

Code	ID #	Location	Hospital	Diagnosis	Specimen	D O C	D O B	Age (y.m.)	Sex	Race	Serotype	pU082 "b"
WA1	88:204	Metropolitan Perth	Princess Margaret	meningitis	csf	Sep-88	11-Oct-87	0.11	nk	A	b	positive
WA2	89:023	S E	Princess Margaret	meningitis	csf	Jan-88	22-May-88	0.8	nk	A	b	positive
WA3	89:139	S E	Princess Margaret	meningitis	csf	May-89	19-Feb-88	1.3	nk	A	b	positive
WA4	94:2	Metropolitan Perth	Princess Margaret	meningitis	csf	Jan-90	31-Aug-89	0.5	nk	A	b	positive
WA5	94:5	Midlands	Princess Margaret	meningitis	csf	Mar-90	23-Jun-89	0.9	nk	A	b	positive
WA6	94:6	S E	Princess Margaret	meningitis	csf	Aug-90	4-Apr-90	0.4	nk	A	b	positive
WA7	94:3	Central	Princess Margaret	meningitis	csf	Dec-90	18-Apr-90	0.8	nk	A	b	positive
WA8	94:1	Metropolitan Perth	Princess Margaret	meningitis	csf	Feb-91	1-Aug-90	0.7	nk	A	b	positive
WA9	84:50	Kimberly	Princess Margaret	meningitis	csf	Apr-84	7-Nov-83	0.6	nk	A	b	positive
WA10	84:72	Central	Princess Margaret	meningitis	csf	May-84	14=Nov-83	0.7	nk	A	b	positive
WA11	non-viable	lyophilized sample										
WA12	85:34	Metropolitan Perth	Princess Margaret	meningitis	csf	Aug-85	23-Mar-85	0.5	nk	A	b	positive
WA13	86:176	Metropolitan Perth	Princess Margaret	meningitis	csf/bld	Oct-86	23-Apr-86	0.6	nk	A	b	positive
WA14	87:57	Metropolitan Perth	Princess Margaret	meningitis	csf/bld	Apr-87	4-Nov-86	0.6	nk	N	b	positive
WA15	87:61	Metropolitan Perth	Princess Margaret	meningitis	csf/bld	May-87	26-Jul-86	10 months	nk	N	b	positive
WA16	87:125	Metropolitan Perth	Princess Margaret	meningitis	csf/bld	Sep-87	23-Jun-86	1.3	nk	N	b	positive
WA17	88:11	Metropolitan Perth	Princess Margaret	meningitis	csf/bld	Jan-88	25-Sep-86	1.4	nk	N	b	positive
WA18	88:207	Metropolitan Perth	Princess Margaret	epiglottitis	bld	Sep-88	3-Feb-85	3.8	nk	N	b	positive
WA19	88:66	Metropolitan Perth	Princess Margaret	epiglottitis	bld	Apr-88	15-Jan-86	2.4	nk	N	b	positive
WA20	87:163	Metropolitan Perth	Princess Margaret	epiglottitis	bld/throat	Oct-87	16-Apr-84	3.7	nk	N	b	positive
WA21	88:8	Metropolitan Perth	Princess Margaret	epiglottitis	bld	Jan-88	5-Dec-86	1.2	nk	N	b	positive
WA22	88:216	Metropolitan Perth	Princess Margaret	meningitis	csf	Oct-88	8=Apr-87	1.6	nk	N	b	positive
WA23	87:172	Metropolitan Perth	Princess Margaret	other	bld	Oct-87	3-Apr-87	1.6	nk	N	b	positive
WA24	87:189	Metropolitan Perth	Princess Margaret	meningitis	csf	Nov-97	16-Dec-86	1.1	nk	N	b	positive
WA25	88:180	Metropolitan Perth	Princess Margaret	meningitis	bld	Sep-88	24-Jun-84	4.3	nk	N	b	positive
WA26	87:195	Metropolitan Perth	Princess Margaret	epiglottitis	bld	Dec-87	26-Jul-84	3.5	nk	N	b	positive
WA27	88:117	Metropolitan Perth	Princess Margaret	meningitis	bld/csf	May-88	6-Jul-86	1.1	nk	N	b	positive
WA28	87:101	Metropolitan Perth	Princess Margaret	epiglottitis	bld	Jun-87	14-Jul-81	5.11	nk	N	b	positive
WA29	87:121	Rural	Princess Margaret	meningitis	csf	Aug-87	4-Feb-87	0.7	nk	N	b	positive
WA30	84:79	Rural	Princess Margaret	meningitis	csf	May-84	4-Oct-83	0.6	nk	N	b	positive

Appendix A.4 (page 2 of 2): Hib Isolates Recovered from Metropolitan Perth and Rural Western Australia

Code	ID #	Location	Hospital	Diagnosis	Specimen	D O C	D O B	Age (y.m.)	Sex	Race	Serotype	pU082 "b"
WA31	84:113	Metropolitan Perth	Princess Margaret	meningitis	bld/csf	Aug-84	27-Feb-84	0.6	nk	N	b	positive
WA32	87:62	Metropolitan Perth	Princess Margaret	other	bld/eye swab	May-87	1-Jul-86	0.11	nk	N	b	positive
WA33	87:122	Metropolitan Perth	Princess Margaret	other	bld	Aug-87	24-Aug-81	6	nk	N	b	positive
WA34	87:106	Metropolitan Perth	Princess Margaret	meningitis	bld	Jul-87	12-Aug-86	0.11	nk	N	b	positive
WA35	87:78	Metropolitan Perth	Princess Margaret	epiglottitis	bld	May-87	4-Sep-80	6.8	nk	N	b	positive
WA36	86:31	Rural	Princess Margaret	epiglottitis	bld	Mar-86	7-Jul-84	1.8	nk	N	b	positive
WA37	87:206	Rural	Princess Margaret	other	bld	Dec-87	9-N0v-87	0.1	nk	N	b	positive
WA38	87:196	Metropolitan Perth	Princess Margaret	other	bld	Nov-87	1-Sep-81	6.2	nk	N	b	positive
WA39	87:197	Metropolitan Perth	Princess Margaret	other	bld/eye swab	Dec-87	31-Jul-87	0.5	nk	N	b	positive
WA40	87:198	Metropolitan Perth	Princess Margaret	epiglottitis	bld	Nov-87	24-Mar-85	2.8	nk	N	b	positive
WA41	87:168	Rural	Princess Margaret	meningitis	bld/csf	Oct-87	30-Aug-84	3.2	nk	N	b	positive
WA42	86:163	Rural	Princess Margaret	epiglottitis	bld	Aug-86	15-Feb-85	1.6	nk	N	b	positive
WA43	84:35	Metropolitan Perth	Princess Margaret	meningitis	csf	Feb-84	27-Oct-82	1.4	nk	N	b	positive
WA44	86:123	Rural	Princess Margaret	meningitis	bld/csf	Jul-86	10-Dec-83	2.7	nk	N	b	positive
WA45	88:170	Rural	Princess Margaret	epiglotittis	bld	Aug-88	16-Oct-85	2.1	nk	N	b	positive
WA46	88:38	Rural	Princess Margaret	epiglotittis	bld	Feb-88	14-Jun-86	1.8	nk	N	b	positive
WA47	87:106	Metropolitan Perth	Princess Margaret	meningitis	bld	Jul-87	12-Aug-86	0.11	nk	N	b	positive

Appendix A.5 (page 1 of 2): Hib Isolates Recovered from the Alice Springs Region

Code	ID #	Location	Hospital	Diagnosis	Specimen	D O C	D O B	Age(y.m.)	Sex	Race	Serotype	pU082 "b"
A1	Q1	Ngukurr	ASH	meningitis	bld	16-Apr-93	25-Nov-92	0.5	m	A	b	positive
A2	Q2	Santa Teresa	ASH	meningitis	bld	8-Apr-93	19-Dec-92	0.4	m	A	b	positive
A3	Q3	Kintore	ASH	pneumonia	bld	27-Feb-93	3-Mar-92	1	f	A	b	positive
A4	Q4	Alice Springs	ASH	cellulitis	bld	8-Feb-93	2-Sep-92	0.5	m	A	b	positive
A5	Q5	Ernabella	ASH	pneumonia	bld	28-Jan-93	92	<1	m	A	b	positive
A6	Q6	Katherine	ASH	meningitis	csf	13-Jan-93	22-May-92	0.8	f	A	b	positive
A7	Q7	Alice Springs	ASH	meningitis	bld	21-Dec-92	17-Jul-92	0.5	m	A	b	positive
A8	Q8	Papunya	ASH	gastro	bld	20-Dec-92	17-Jun-92	0.6	m	A	b	positive
A9	Q9	Yarralin	ASH	pneumonia	bld	28-Nov-92	2-Jun-92	0.6	f	A	b	positive
A10isNT10	Q10	duplicate										
A11	Q11	Alice Springs region	ASH	well	npa	not known	28-Jul-90	nk	f	A	b	positive
A12	Q12	Alice Springs region	ASH	well	npa	not known	12-Jul-90	nk	m	A	b	positive
A13	Q13	Alice Springs region	ASH	well	npa	not known	21-Feb-92	nk	m	A	b	positive
A14	Q14	Alice Springs region	ASH	well	npa	not known	3-Apr-91	nk	m	A	b	positive
A15	Q15	Alice Springs region	ASH	well	npa	not known	21-Feb-92	nk	m	A	b	positive
A16	Q16	Alice Springs region	ASH	well	npa	not known	4-Mar-91	nk	f	A	b	positive
A17	Q17	Alice Springs region	ASH	ALRI	npa	not known	91-Apr-1	nk	m	A	b	positive
A18	Q18	Alice Springs region	ASH	ALRI	npa	not known	3-Apr-91	nk	m	A	b	positive
A19	Q19	Alice Springs region	ASH	ALRI	npa	not known	2-Dec-90	nk	f	A	b	positive
A20	Q20	Alice Springs region	ASH	ALRI	npa	not known	2-Dec-90	nk		A	b	positive
A21 pair	Q21	Hermannsburg	ASH	meningitis	blood	not known	31-May-92	0.4	f	A	b	positive
A22 pair	Q22	see A21 - diff. RFLP	ASH	meningitis	csf	not known	31-May-92	0.4	f	A	b	positive
A23	Q23	Katherine	ASH	febrile	bld	5-Sep-92	5-Feb-92	0.7	f	A	b	positive
A24 is NT9	Q24	duplicate										
A25 is NTHi	Q25	Hermannsburg	ASH	heart/resp failure	bld	30-Aug-92	20-Dec-90	1.8	f	A	b/c	NEGATIVE
A26	Q26	Alice Springs area	ASH	bronchiolitis	bld	11-Aug-92	6-Mar-92	0.5	m	A	b	positive
A27	Q27	Yambah	ASH	pneumonia	bld	3-Aug-92	11-Feb-92	0.6	m	A	b	positive
A28	Q28	Fregon, SA	ASH	gastro	bld	26-Jul-92	31-Aug-91	1.1	m	A	b	positive
A29 pair	Q29	Tennant Creek	ASH	pneumonia	bld	11-Jul-92	20-Feb-92	0.5	m	A	b	positive
A30 pair	Q30	see A29 - same RFLP	ASH	see A29	pleural fluid	see A29	see A29	see A29	see A29	see A29	b	positive
A31	Q31	Kintore	ASH	pneumonia	bld	9-Jul-92	18-Sep-91	10 mo.	m	A	b	positive

Appendix A.5 (page 2 of 2): Hib Isolates Recovered from the Alice Springs Region

Code	ID #	Location	Hospital	Diagnosis	Specimen	D O C	D O B	Age(y.m.)	Sex	Race	Serotype	pU082 "b"
A32	Q32	Ernabella, SA	ASH	gastro	bld	30-May-92	9-Aug-92	10 months	f	A	b	positive
A33	Q33	Mt. Davis	ASH	gastro	bld	11-May-92	6-May-92	1	m	A	b	positive
A34 pair	Q34	Alice Springs region	ASH	meningitis	bld	13-May-92	12-Dec-91	0.5	f	A	b	positive
A35 pair	Q35	see A34 - diff. RFLP	ASH	see A34	csf	see A34	see A34	see A34	see A34	see A34	b	positive
A36	Q36	Katherine	ASH	meningitis	bld	26-May-92	24-Mar-91	1.2	f	A	b	positive
A37	Q37	Mt. Allen	ASH	meningitis	bld	1-May-92	31-Jan-92	0.3	m	A	b	positive
A38	Q38	Kintore	ASH	gastro	bld	5-Apr-92	16-Jul-91	0.9	m	A	b	positive
A39	Q39	Ali Curang	ASH	failure to thrive	bld	11-Jan-92	8-Dec-90	0.1	f	A	b	positive
A40	Q40	Katherine	ASH	pneumonia	bld	8-Jan-92	17-Feb-91	0.11	f	A	b	positive
A41	5C1	Alice Springs region	ASH	meningitis	not known	8-Jan-89	not known	nk	f	A	b	positive
A42	10C1	Alice Springs region	ASH	eye	not known	17-Jul-85	not known	nk	m	A	b	positive
A43	23C1	Alice Springs region	ASH	gastro	not known	10-May-90	not known	nk	m	A	b	positive
A44	24C1	Alice Springs region	ASH	gastro & pneumo	not known	27-Mar-89	not known	nk	f	A	b	positive
A45	25C1	Alice Springs region	ASH	pneumo	not known	8-May-90	not known	nk	m	A	b	positive
A46	26C1	Alice Springs region	ASH	gastro & ? pneumo	not known	89	not known	nk	f	A	b	positive
A47 pair	27C1	Alice Springs region	ASH	otitis media & men	not known	19-Nov-89	not known	nk	f	A	b	positive
A48 pair	28C1	see A47 - same RFLP	see A47	see A47	see A47	see A47	not known	nk	see A47	A	b	positive
A49	30C1	Alice Springs region	ASH	gastro & septic	not known	8-Jan-89	not known	nk	m	A	b	positive
A50	31C1	Alice Springs region	ASH	gastro & entero	not known	7-Apr-90	not known	nk	m	A	b	positive
NT1		ASH - tourist	ASH	meningitis	csf	7-Jul-89	4-Nov-88	0.8	m	N	b	positive
NT2		Alice Springs region	ASH	meningitis	bld/csf	12-Jan-91	19-Nov-89	0.2	f	N	b	positive
NT3isNT2		Alice Springs region	ASH	meningitis	bld/csf	see NT2	see NT2	see NT2	as above	N	b	positive
NT4		Alice Springs region	ASH	meningitis	bld/csf	1-Apr-91	4-Nov-88	2.5	f	N	b	positive
NT5isNT4		Alice Springs region	ASH	meningitis	bld/csf	see NT4	see NT4	see NT4	as above	N	b	positive
NT6		Tourist from Sydney	ASH	invasive	bld	11-Apr-91	14-Jan-54	37.3	f	N	b	positive
NT7		Alice Springs region	ASH	invasive	bld	15-Aug-91	5-Oct-90	10 mo.	m	N	b	positive
NT8		Alice Springs region	ASH	invasive	bld	21-Aug-91	4-Dec-89	1.9	m	N	b	positive
NT9		Alice Springs region	ASH	invasive	bld	2-Sep-92	17-Oct-90	1.11	m	N	b	positive
NT10		Katherine	ASH	invasive	bld	7-Oct-92	30-Aug-90	2.1	f	N	b	positive
NT11		Katherine	ASH	invasive	bld	13-Nov-92	Feb-92	0.9	m	N	b	positive
NT12		Katherine	ASH	invasive	bld	21-May-93	20-Feb-47	46.3	f	N	b	positive

Appendix A.6: Hib Isolates Recovered from Bathurst Island

Code	ID #	Location	Hospital	Diagnosis	Specimen	D O C	D O B	Age(y.m.)	Sex	Race	Serotype	Biotype	pU082 "b"
B1	909-1	Nguiu, Bathhurst Is.	M S H R	carrier	nasopharynx	24-Mar-92	not known	0.2	nk	A	b	II	positive
B2	911-1	Nguiu, Bathhurst Is.	M S H R	carrier	nasopharynx	14-Apr-92	not known	0.2	nk	A	b	III	positive
B3	911-2	Nguiu, Bathhurst Is.	M S H R	carrier	nasopharynx	14-Apr-92	not known	0.2	nk	A	b	IV	positive
B4	911-4	Nguiu, Bathhurst Is.	M S H R	carrier	nasopharynx	14-Apr-92	not known	0.2	nk	A	b	II	positive
B5	913-1	Nguiu, Bathhurst Is.	M S H R	carrier	nasopharynx	20-May-92	not known	0.3	nk	A	b	I	positive
B6	914-1	Nguiu, Bathhurst Is.	M S H R	carrier	nasopharynx	21-Jul-92	not known	0.4	nk	A	b	III	positive
B7	914-3	Nguiu, Bathhurst Is.	M S H R	carrier	nasopharynx	2-Dec-92	not known	0.8	nk	A	b	I	positive
B8	923-1	Nguiu, Bathhurst Is.	M S H R	carrier	nasopharynx	21-Jul-92	not known	0.1	nk	A	b	I	positive
B9	non-viable	on receipt											
B10	925-1	Nguiu, Bathhurst Is.	M S H R	carrier	nasopharynx	21-Jul-92	not known	0.08	nk	A	b	I	positive
B11	926-2	Nguiu, Bathhurst Is.	M S H R	carrier	nasopharynx	1-Dec-92	not known	0.4	nk	A	b	I	positive
B12	928-1	Nguiu, Bathhurst Is.	M S H R	carrier	nasopharynx	5-Jan-93	not known	0.5	nk	A	b	I	positive
B13	931-1	Nguiu, Bathhurst Is.	M S H R	otitis media?	ear discharge	28-Apr-93	not known	0.7	nk	A	b	II	positive
B14	934-1	Nguiu, Bathhurst Is.	M S H R	carrier	nasopharynx	2-Dec-92	not known	0.1	nk	A	b	I	positive
B15	934-3	Nguiu, Bathhurst Is.	M S H R	carrier	nasopharynx	10-Feb-93	not known	0.3	nk	A	b	I	positive
B16	935-1	Nguiu, Bathhurst Is.	M S H R	carrier	nasopharynx	3-Mar-93	not known	0.4	nk	A	b	I	positive
B17	non-viable	on receipt											
B18	943-3	Nguiu, Bathhurst Is.	M S H R	carrier	nasopharynx	10-Aug-93	not known	0.3	nk	A	b	I	positive
B19	943-2	Nguiu, Bathhurst Is.	M S H R	carrier	nasopharynx	20-Sep-93	not known	0.4	nk	A	b	II	positive
B20	947-1	Nguiu, Bathhurst Is.	M S H R	carrier	nasopharynx	5-Oct-93	not known	0.4	nk	A	b	I	positive
B21	950-1	Nguiu, Bathhurst Is.	M S H R	carrier	nasopharynx	21-Sep-93	not known	0.3	nk	A	b	I	positive

APPENDIX B: Table 3.7 Distance matrix of the 1-*F* values* of 16 *Sma*I-RFLP types found among 20 meningitis isolates

M1	0.0000														
M2	0.1429	0.0000													
M3[a]	0.2143	0.2143	0.0000												
M5[b]	0.1852	0.1111	0.1111	0.0000											
M6	0.2414	0.2414	0.1034	0.1429	0.0000										
M7	0.6000	0.5333	0.6000	0.5862	0.5484	0.0000									
M8	0.2414	0.1724	0.1724	0.0714	0.1333	0.5484	0.0000								
M9	0.6552	0.5862	0.5862	0.5714	0.5333	0.3548	0.5333	0.0000							
M10	0.2414	0.2414	0.0345	0.1429	0.1333	0.6129	0.2000	0.6000	0.0000						
M11	0.2414	0.1724	0.1034	0.0714	0.2000	0.6129	0.1333	0.6000	0.1333	0.0000					
M13	0.7241	0.6552	0.7241	0.6429	0.6667	0.8065	0.6000	0.6667	0.7333	0.6667	0.0000				
M14	0.3333	0.3333	0.2667	0.2414	0.2258	0.6250	0.2258	0.4839	0.2903	0.2903	0.5484	0.0000			
M15	0.3571	0.2857	0.2857	0.2593	0.2414	0.5333	0.2414	0.5172	0.3103	0.3103	0.6552	0.2667	0.0000		
M17	0.2593	0.1852	0.1852	0.0769	0.2143	0.6552	0.1429	0.5714	0.2143	0.1429	0.6429	0.3103	0.3333	0.0000	
M19	0.2143	0.2143	0.0714	0.1111	0.0345	0.6000	0.1724	0.5862	0.1034	0.1724	0.7241	0.2667	0.2857	0.1852	0.0000
M20	0.0714	0.2143	0.2143	0.1852	0.2414	0.6000	0.2414	0.6552	0.2414	0.2414	0.7241	0.3333	0.3571	0.2593	0.2143
	M1	M2	M3[a]	M5[b]	M6	M7	M8	M9	M10	M11	M13	M14	M15	M17	M19

The lightly shaded numbers represent the *F* values ≥0.9 (1-the shaded value) that indicate clonally related pairs of isolates that have a one to three fragment difference. The boxed number is the 1-*F* value of the most diverse pair of isolates, M13–M7, that had an *F* value of 0.19. This pair shared only 3 of the 28 different fragments found between them.

*The *F* value is 1 minus the value found in this matrix because the RAPDistance Package program used for constructing the matrix has been designed to subtract the *F* value from 1. Therefore, the lowest values in this matrix represent the most closely related pairs of isolates and the highest values represent the most diverse pairs of isolates. The zero values shown in this matrix are the values relating to the comparison of identical pairs that would yield an *F* value of 1 represented by a 0.0000 value in a matrix generated using the RAPDistance Package.

[a]*Sma*I-M3 included isolates M3 and M4.
[b]*Sma*I-M5 included isolates M5, M12, M16, and M18.

APPENDIX B: Table 3.8 Distance matrix of the 1-*F* values* of 12 *Sma*I-RFLP types found among 14 epiglottitis isolates

E1	0.0000											
E2[a]	0.3333	0.0000										
E3	0.3548	0.1724	0.0000									
E4	0.4000	0.2143	0.1034	0.0000								
E5	0.4000	0.4286	0.4483	0.5000	0.0000							
E6	0.3750	0.1333	0.1613	0.2000	0.4667	0.0000						
E7	0.3548	0.1034	0.2000	0.2414	0.3793	0.2258	0.0000					
E9[b]	0.3793	0.1111	0.2143	0.1852	0.4815	0.2414	0.0714	0.0000				
E11	0.3793	0.3333	0.3571	0.4074	0.1111	0.3793	0.2857	0.3846	0.0000			
E12	0.3103	0.0370	0.1429	0.1852	0.4074	0.1724	0.0714	0.0769	0.3077	0.0000		
E13	0.3548	0.1724	0.1333	0.1724	0.4483	0.0968	0.2000	0.2143	0.3571	0.1429	0.0000	
E15	0.4194	0.1724	0.2667	0.3103	0.5172	0.2258	0.2667	0.2857	0.4286	0.2143	0.2667	0.0000
	E1	E2[a]	E3	E4	E5	E6	E7	E9[b]	E11	E12	E13	E15

The lightly shaded numbers represent the *F* values ≥0.9 (1-the shaded value) that indicate clonally related pairs of isolates that have a one to three fragment difference. The boxed number is the 1-*F* value of the most diverse pair of isolates, E5–E15, that had an *F* value of 0.48. This pair shared only 7 of the 22 different fragments found between them.

*The *F* value is 1 minus the value found in this matrix because the RAPDistance Package program used for constructing the matrix has been designed to subtract the *F* value from 1. Therefore, the lowest values in this matrix represent the most closely related pairs of isolates and the highest values represent the most diverse pairs of isolates. The zero values shown in this matrix are the values relating to the comparison of identical pairs that would yield an *F* value of 1 represented by a 0.0000 value in a matrix generated using the RAPDistance Package.

[a]*Sma*I-E2 included isolates E2 and E14.
[b]*Sma*I-E9 included isolates E9 and E10.

APPENDIX B: Table 3.9 Distance matrix of the 1-*F* values* of 26 *Sma*I-RFLP types found among 34 meningitis and epiglottitis isolates

M1	0.0000													
M2	0.1429	0.0000												
M3[a]	0.2143	0.2143	0.0000											
M5[b]	0.1852	0.1111	0.1111	0.0000										
M6	0.2414	0.2414	0.1034	0.1429	0.0000									
M7	0.6000	0.5333	0.6000	0.5862	0.5484	0.0000								
M8[c]	0.2414	0.1724	0.1724	0.0714	0.1333	0.5484	0.0000							
M9	0.6552	0.5862	0.5862	0.5714	0.5333	0.3548	0.5333	0.0000						
M10	0.2414	0.2414	0.0345	0.1429	0.1333	0.6129	0.2000	0.6000	0.0000					
M11	0.2414	0.1724	0.1034	0.0714	0.2000	0.6129	0.1333	0.6000	0.1333	0.0000				
M13	0.7241	0.6552	0.7241	0.6429	0.6667	0.8065	0.6000	0.6667	0.7333	0.6667	0.0000			
M14	0.3333	0.3333	0.2667	0.2414	0.2258	0.6250	0.2258	0.4839	0.2903	0.2903	0.5484	0.0000		
M15	0.3571	0.2857	0.2857	0.2593	0.2414	0.5333	0.2414	0.5172	0.3103	0.3103	0.6552	0.2667	0.0000	
M17	0.2593	0.1852	0.1852	0.0769	0.2143	0.6552	0.1429	0.5714	0.2143	0.1429	0.6429	0.3103	0.3333	0.0000
M19	0.2143	0.2143	0.0714	0.1111	0.0345	0.6000	0.1724	0.5862	0.1034	0.1724	0.7241	0.2667	0.2857	0.1852
M20	0.0714	0.2143	0.2143	0.1852	0.2414	0.6000	0.2414	0.6552	0.2414	0.2414	0.7241	0.3333	0.3571	0.2593
E1	0.4000	0.3333	0.3333	0.3103	0.3548	0.6250	0.3548	0.6129	0.3548	0.3548	0.5484	0.4375	0.3333	0.2414
E2[d]	0.2143	0.1429	0.1429	0.0370	0.1724	0.5333	0.1034	0.5172	0.1724	0.1034	0.6552	0.2667	0.2857	0.1111
E3	0.2414	0.2414	0.0345	0.1429	0.1333	0.5484	0.2000	0.5333	0.0667	0.1333	0.7333	0.2903	0.3103	0.2143
E4	0.2857	0.2857	0.1429	0.1852	0.1724	0.5333	0.2414	0.5172	0.1724	0.2414	0.7931	0.3333	0.3571	0.2593
E5	0.5000	0.4286	0.4286	0.4074	0.3793	0.6000	0.3793	0.5172	0.4483	0.4483	0.5862	0.4000	0.1429	0.4074
E6	0.2667	0.2667	0.1333	0.1724	0.0968	0.5625	0.2258	0.5484	0.1613	0.2258	0.7419	0.3125	0.3333	0.2414
E9[e]	0.2593	0.1852	0.1852	0.0769	0.2143	0.5862	0.0714	0.5714	0.2143	0.1429	0.7143	0.3103	0.3333	0.1538
E11	0.4074	0.3333	0.3333	0.3077	0.2857	0.5172	0.2857	0.5000	0.3571	0.3571	0.6429	0.3103	0.0370	0.3846
E13	0.2414	0.2414	0.1034	0.1429	0.0667	0.6129	0.2000	0.6000	0.1333	0.2000	0.7333	0.2903	0.3103	0.2143
E15	0.1034	0.2414	0.2414	0.2143	0.2667	0.4839	0.2667	0.6000	0.2667	0.2667	0.7333	0.3548	0.3793	0.2857
	M1	M2	M3[a]	M5[b]	M6	M7	M8[c]	M9	M10	M11	M13	M14	M15	M17

M20	0.2143	0.0000										
E1	0.3333	0.4000	0.0000									
E2[d]	0.1429	0.2143	0.3333	0.0000								
E3	0.1034	0.2414	0.3548	0.1724	0.0000							
E4	0.1429	0.2857	0.4000	0.2143	0.1034	0.0000						
E5	0.4286	0.5000	0.4000	0.4286	0.4483	0.5000	0.0000					
E6	0.0667	0.2667	0.3750	0.1333	0.1613	0.2000	0.4667	0.0000				
E9[e]	0.1852	0.2593	0.3793	0.1111	0.2143	0.1852	0.4815	0.2414	0.0000			
E11	0.3333	0.4074	0.3793	0.3333	0.3571	0.4074	0.1111	0.3793	0.3846	0.0000		
E13	0.0345	0.2414	0.3548	0.1724	0.1333	0.1724	0.4483	0.0968	0.2143	0.3571	0.0000	
E15	0.2414	0.1034	0.4194	0.1724	0.2667	0.3103	0.5172	0.2258	0.2857	0.4286	0.2667	0.0000
	M19	M20	E1	E2[d]	E3	E4	E5	E6	E9[e]	E11	E13	E15

The lightly shaded numbers represent the F values ≥0.9 (1-the shaded value) that indicate clonally related pairs of isolates that have a one to three fragment difference. The boxed number is the 1-F value of the most diverse pair of isolates, M13–M7, that had an F value of 0.19. This pair shared only 3 of the 28 different fragments found between them.

*The F value is 1 minus the value found in this matrix because the RAPDistance Package program used for constructing the matrix has been designed to subtract the F value from 1. Therefore, the lowest values in this matrix represent the most closely related pairs of isolates and the highest values represent the most diverse pairs of isolates. The zero values shown in this matrix are the values relating to the comparison of identical pairs that would yield an F value of 1 represented by a 0 value in a matrix generated using the RAPDistance Package.

[a]*Sma*I-M3 included isolates M3 and M4.
[b]*Sma*I-M5 included isolates M5, M12, M16, M18, and E12.
[c]*Sma*I-M8 included isolates M8 and E7.
[d]*Sma*I-E2 included isolates E2 and E14.
[e]*Sma*I-E9 included isolates E9 and E10.

APPENDIX B: Table 3.10 Distance matrix of the 1-*F* values* of 15 *Apa*I-RFLP types found among 20 meningitis isolates

	M1	M2	M3	M4	M5[a]	M7	M8	M9	M10	M13	M14	M15	M17	M19
M1	0.0000													
M2	0.0857	0.0000												
M3	0.1351	0.2222	0.0000											
M4	0.1892	0.2222	0.1579	0.0000										
M5[a]	0.1429	0.1176	0.1667	0.1111	0.0000									
M7	0.3143	0.2941	0.3333	0.3333	0.2353	0.0000								
M8	0.2000	0.1765	0.2222	0.1667	0.0588	0.1765	0.0000							
M9	0.3158	0.2973	0.3333	0.2821	0.1892	0.4054	0.2432	0.0000						
M10	0.1351	0.2222	0.0526	0.1053	0.1111	0.3333	0.1667	0.2821	0.0000					
M13	0.5676	0.5556	0.5263	0.4211	0.4444	0.4444	0.4444	0.3846	0.4737	0.0000				
M14	0.3846	0.3684	0.3500	0.3000	0.3158	0.4737	0.3158	0.1220	0.3000	0.4000	0.0000			
M15	0.3514	0.3333	0.3158	0.2105	0.2222	0.2778	0.2778	0.3333	0.2632	0.3684	0.4000	0.0000		
M17	0.2000	0.1765	0.2222	0.1667	0.0588	0.2353	0.1176	0.2432	0.1667	0.4444	0.3684	0.2778	0.0000	
M19	0.1667	0.1429	0.1351	0.1351	0.0286	0.2571	0.0857	0.2105	0.0811	0.4595	0.3333	0.2432	0.0857	0.0000
M20	0.0556	0.0857	0.1351	0.1892	0.1429	0.3143	0.2000	0.3158	0.1351	0.5676	0.3846	0.3514	0.2000	0.1111

The lightly shaded numbers represent the *F* values ≥0.9 (1-the shaded value) that indicate clonally or closely related pairs of isolates that have a one to three fragment difference. The boxed number is the 1-*F* value of the most diverse pairs of isolates, M1 –M13 and M13–M20, that had an *F* value of 0.43. These pairs each shared only 8 of the 29 different fragments found between them.

*The *F* value is 1 minus the value found in this matrix because the RAPDistance Package program used for constructing the matrix has been designed to subtract the *F* value from 1. Therefore, the lowest values in this matrix represent the most closely related pairs of isolates and the highest values represent the most diverse pairs of isolates. The zero values shown in this matrix are the values relating to the comparison of identical pairs that would yield an *F* value of 1 represented by a 0 value in a matrix generated using the RAPDistance Package.

[a]*Apa*I-M5 included isolates M5, M6, M11, M12, M16, and M18.

APPENDIX B: Table 3.11 Distance matrix of the 1-*F* values* of 9 *Apa*I-RFLP types found among 14 epiglottitis isolates

E1	0.0000								
E2	0.2727	0.0000							
E3	0.3714	0.1667	0.0000						
E4	0.3529	0.1429	0.1351	0.0000					
E5[a]	0.2000	0.1667	0.2632	0.1892	0.0000				
E6[b]	0.3333	0.1176	0.1111	0.0286	0.2222	0.0000			
E7	0.3333	0.1765	0.1667	0.0857	0.2778	0.0588	0.0000		
E13	0.3529	0.1429	0.1351	0.0556	0.2432	0.0286	0.0857	0.0000	
E15	0.4706	0.2571	0.1892	0.1667	0.3514	0.1429	0.2000	0.1111	0.0000
	E1	E2	E3	E4	E5[a]	E6[b]	E7	E13	E15

The lightly shaded numbers represent the *F* values ≥0.9 (1-the shaded value)that indicate clonally or closely related pairs of isolates that have a one to three fragment difference. The boxed number is the 1-*F* value of the most diverse pair of isolates, E1 –E15, that had an *F* value of 0.53. This pair shared only 9 of the 25 different fragments found between them.

*The *F* value is 1 minus the value found in this matrix because the RAPDistance Package program used for constructing the matrix has been designed to subtract the *F* value from 1. Therefore, the lowest values in this matrix represent the most closely related pairs of isolates and the highest values represent the most diverse pairs of isolates. The zero values shown in this matrix are the values relating to the comparison of identical pairs that would yield an *F* value of 1 represented by a 0 value in a matrix generated using the RAPDistance Package.

[a] *Apa*I-E5 included isolates E5 and E11.
[b] *Apa*I-E6 included isolates E6, E9, E10, E12, and E14.

APPENDIX B: Table 3.12 Distance matrix of the 1-*F* values* of 20 *Apa*I-RFLP types** found among 34 meningitis and epiglottitis isolates

M1	0.0000														
M2	0.0857	0.0000													
M3	0.1351	0.2222	0.0000												
M4	0.1892	0.2222	0.1579	0.0000											
M5[a]	0.1429	0.1176	0.1667	0.1111	0.0000										
M7	0.3143	0.2941	0.3333	0.3333	0.2353	0.0000									
M8[b]	0.2000	0.1765	0.2222	0.1667	0.0588	0.1765	0.0000								
M9	0.3158	0.2973	0.3333	0.2821	0.1892	0.4054	0.2432	0.0000							
M10	0.1351	0.2222	0.0526	0.1053	0.1111	0.3333	0.1667	0.2821	0.0000						
M13	0.5676	0.5556	0.5263	0.4211	0.4444	0.4444	0.4444	0.3846	0.4737	0.0000					
M14	0.3846	0.3684	0.3500	0.3000	0.3158	0.4737	0.3158	0.1220	0.3000	0.4000	0.0000				
M15[c]	0.3514	0.3333	0.3158	0.2105	0.2222	0.2778	0.2778	0.3333	0.2632	0.3684	0.4000	0.0000			
M17	0.2000	0.1765	0.2222	0.1667	0.0588	0.2353	0.1176	0.2432	0.1667	0.4444	0.3684	0.2778	0.0000		
M19	0.1667	0.1429	0.1351	0.1351	0.0286	0.2571	0.0857	0.2105	0.0811	0.4595	0.3333	0.2432	0.0857	0.0000	
M20[d]	0.0556	0.0857	0.1351	0.1892	0.1429	0.3143	0.2000	0.3158	0.1351	0.5676	0.3846	0.3514	0.2000	0.1111	0.0000
E1	0.4706	0.4545	0.4286	0.3143	0.3333	0.3333	0.3333	0.4444	0.3714	0.4286	0.4595	0.2000	0.2727	0.3529	0.4706
E2	0.2571	0.2353	0.2222	0.1111	0.1176	0.3529	0.1765	0.2973	0.1667	0.5000	0.3684	0.1667	0.1765	0.1429	0.2571
E3	0.1351	0.2222	0.1053	0.1053	0.1111	0.3333	0.1667	0.2821	0.0526	0.4737	0.3000	0.2632	0.1667	0.1351	0.1892
E4	0.1667	0.1429	0.1892	0.1351	0.0286	0.2000	0.0857	0.2105	0.1351	0.4054	0.3333	0.1892	0.0857	0.0556	0.1667
E13	0.1111	0.0857	0.1892	0.1351	0.0286	0.2571	0.0857	0.2105	0.1351	0.4595	0.3333	0.2432	0.0857	0.0556	0.1111
E15[d]	0.0556	0.0857	0.1351	0.1892	0.1429	0.3143	0.2000	0.3158	0.1351	0.5676	0.3846	0.3514	0.2000	0.1111	0.0000
	M1	M2	M3	M4	M5[a]	M7	M8[b]	M9	M10	M13	M14	M15[c]	M17	M19	M20[d]

E1	0.0000					
E2	0.2727	0.0000				
E3	0.3714	0.1667	0.0000			
E4	0.3529	0.1429	0.1351	0.0000		
E13	0.3529	0.1429	0.1351	0.0556	0.0000	
E15[d]	0.4706	0.2571	0.1892	0.1667	0.1111	0.0000
	E1	E2	E3	E4	E13	E15[d]

The lightly shaded numbers represent the *F* values ≥0.9 (1-the shaded value) that indicate clonally or closely related pairs of isolates that have a one to three fragment difference. The boxed number is the 1-*F* value of the most diverse pairs of isolates, M1–M13 and M13–M20, that had an *F* value of 0.43. These pairs each shared only 8 of the 29 different fragments found between them.

*The *F* value is 1 minus the value found in this matrix because the RAPDistance Package program used for constructing the matrix has been designed to subtract the *F* value from 1. Therefore, the lowest values in this matrix represent the most closely related pairs of isolates and the highest values represent the most diverse pairs of isolates. The zero values shown in this matrix are the values relating to the comparison of identical pairs that would yield an *F* value of 1 represented by a 0 value in a matrix generated using the RAPDistance Package.

[a] *Apa*I-M5 included isolates M5, M6, M11, M12, M16, M18, E6, E9, E10, E12, and E14.
[b] *Apa*I-M8 included isolates M8 and E7.
[c] *Apa*I-M15 included isolates M15, E5, and E11.
[d] *Apa*I-M20 and included isolates M20 and E15. [Both M20 and E15 are included in the matrix rather than only one representative isolate per type because the NJTREE program would not compute if one of the isolates was omitted.]

APPENDIX B: Table 3.13 Distance matrix of the 1-*F* values* of 17 *Sma*I/*Apa*I-RFLP types found among 20 meningitis isolates

	M1	M2	M3	M4	M5[a]	M6	M7	M8	M9	M10	M11	M13	M14	M15	M17	M19
M1	0.0000															
M2	0.1111	0.0000														
M3	0.1692	0.2188	0.0000													
M4	0.2000	0.2188	0.0909	0.0000												
M5[a]	0.1613	0.1148	0.1429	0.1111	0.0000											
M6	0.1875	0.1746	0.1385	0.1077	0.0645	0.0000										
M7	0.4462	0.4063	0.4545	0.4545	0.3968	0.3846	0.0000									
M8	0.2188	0.1746	0.2000	0.1692	0.0645	0.0938	0.3538	0.0000								
M9	0.4627	0.4242	0.4412	0.4118	0.3538	0.3433	0.3824	0.3731	0.0000							
M10	0.1818	0.2308	0.0448	0.0746	0.1250	0.1212	0.4627	0.1818	0.4203	0.0000						
M11	0.1875	0.1429	0.1385	0.1077	0.0323	0.0938	0.4154	0.0938	0.3731	0.1212	0.0000					
M13	0.6364	0.6000	0.6119	0.5522	0.5313	0.5455	0.6119	0.5152	0.5072	0.5882	0.5455	0.0000				
M14	0.3623	0.3529	0.3143	0.2857	0.2836	0.2754	0.5429	0.2754	0.2778	0.2958	0.3043	0.4648	0.0000			
M15	0.3538	0.3125	0.3030	0.2424	0.2381	0.2308	0.3939	0.2615	0.4118	0.2836	0.2615	0.4925	0.3429	0.0000		
M17	0.2258	0.1803	0.2063	0.1746	0.0667	0.1290	0.4286	0.1290	0.3846	0.1875	0.0968	0.5313	0.3433	0.3016	0.0000	
M19	0.1875	0.1746	0.1077	0.1077	0.0645	0.0313	0.4154	0.1250	0.3731	0.0909	0.0938	0.5758	0.3043	0.2615	0.1290	0.0000
M20	0.0625	0.1429	0.1692	0.2000	0.1613	0.1875	0.4462	0.2188	0.4627	0.1818	0.1875	0.6364	0.3623	0.3538	0.2258	0.1563

The lightly shaded numbers represent the *F* values ≥0.9 (1-the shaded value) that indicate clonally related pairs of isolates that have a one to three fragment difference. The boxed number is the 1- *F* value of the most diverse pair of isolates, M13 –M20, that had an *F* value of 0.36. This pair shared only 12 of the 54 different fragments found between them.

*The *F* value is 1 minus the value found in this matrix because the RAPDistance Package program used for constructing the matrix has been designed to subtract the *F* value from 1 Therefore, the lowest values in this matrix represent the most closely related pairs of isolates and the highest values represent the most diverse pairs of isolates. The zero values shown in this matrix are the values relating to the comparison of identical pairs that would yield an *F* value of 1 represented by a 0 value in a matrix generated using the RAPDistance Package.

[a]*Sma*I/*Apa*I-M5 included isolates M5, M12, M16, and M18.

APPENDIX B: Table 3.14 Distance matrix of the 1-*F* values* of 13 *Sma*I/*Apa*I-RFLP types found among 14 epiglottitis isolates

E1	0.0000												
E2	0.3016	0.0000											
E3	0.3636	0.1692	0.0000										
E4	0.3750	0.1746	0.1212	0.0000									
E5	0.2923	0.2813	0.3433	0.3231	0.0000								
E6	0.3538	0.1250	0.1343	0.1077	0.3333	0.0000							
E7	0.3438	0.1429	0.1818	0.1563	0.3231	0.1385	0.0000						
E9[a]	0.3548	0.1148	0.1563	0.0968	0.3333	0.1111	0.0645	0.0000					
E11	0.2813	0.2381	0.3030	0.2813	0.0462	0.2923	0.2813	0.2903	0.0000				
E12	0.3226	0.0820	0.1250	0.0968	0.3016	0.0794	0.0645	0.0333	0.2581	0.0000			
E13	0.3538	0.1563	0.1343	0.1077	0.3333	0.0606	0.1385	0.1111	0.2923	0.0794	0.0000		
E14	0.3438	0.0794	0.1515	0.1250	0.3231	0.0769	0.0938	0.0645	0.2813	0.0323	0.1077	0.0000	
E15	0.4462	0.2188	0.2239	0.2308	0.4242	0.1818	0.2308	0.2063	0.3846	0.1746	0.2121	0.1692	0.0000
	E1	E2	E3	E4	E5	E6	E7	E9[a]	E11	E12	E13	E14	E15

The lightly shaded numbers represent the *F* values ≥0.9 (1-the shaded value) that indicate clonally related pairs of isolates that have a one to three fragment difference. The boxed number is the 1- *F* value of the most diverse pair of isolates, E1 –E15, that had an *F* value of 0.56. This pair shared only 18 of the 65 different fragments found between them.

*The *F* value is 1 minus the value found in this matrix because the RAPDistance Package program used for constructing the matrix has been designed to subtract the *F* value from 1. Therefore, the lowest values in this matrix represent the most closely related pairs of isolates and the highest values represent the most diverse pairs of isolates. The zero values shown in this matrix are the values relating to the comparison of identical pairs that would yield an *F* value of 1 represented by a 0 value in a matrix generated using the RAPDistance Package.

[a]*Sma*I/*Apa*I-E9 included isolates E9 and E10.

APPENDIX B: Table 3.15 Distance matrix of the 1-*F* values* of 28 *Sma*I/*Apa*I-RFLP types found among 34 meningitis and epiglottitis isolates

	M1	M2	M3	M4	M5[a]	M6	M7	M8[b]	M9	M10	M11	M13	M14	M15
M1	0.0000													
M2	0.1111	0.0000												
M3	0.1692	0.2188	0.0000											
M4	0.2000	0.2188	0.0909	0.0000										
M5[a]	0.1613	0.1148	0.1429	0.1111	0.0000									
M6	0.1875	0.1746	0.1385	0.1077	0.0645	0.0000								
M7	0.4462	0.4063	0.4545	0.4545	0.3968	0.3846	0.0000							
M8[b]	0.2188	0.1746	0.2000	0.1692	0.0645	0.0938	0.3538	0.0000						
M9	0.4627	0.4242	0.4412	0.4118	0.3538	0.3433	0.3824	0.3731	0.0000					
M10	0.1818	0.2308	0.0448	0.0746	0.1250	0.1212	0.4627	0.1818	0.4203	0.0000				
M11	0.1875	0.1429	0.1385	0.1077	0.0323	0.0938	0.4154	0.0938	0.3731	0.1212	0.0000			
M13	0.6364	0.6000	0.6119	0.5522	0.5313	0.5455	0.6119	0.5152	0.5072	0.5882	0.5455	0.0000		
M14	0.3623	0.3529	0.3143	0.2857	0.2836	0.2754	0.5429	0.2754	0.2778	0.2958	0.3043	0.4648	0.0000	
M15	0.3538	0.3125	0.3030	0.2424	0.2381	0.2308	0.3939	0.2615	0.4118	0.2836	0.2615	0.4925	0.3429	0.0000
M17	0.2258	0.1803	0.2063	0.1746	0.0667	0.1290	0.4286	0.1290	0.3846	0.1875	0.0968	0.5313	0.3433	0.3016
M19	0.1875	0.1746	0.1077	0.1077	0.0645	0.0313	0.4154	0.1250	0.3731	0.0909	0.0938	0.5758	0.3043	0.2615
M20	0.0625	0.1429	0.1692	0.2000	0.1613	0.1875	0.4462	0.2188	0.4627	0.1818	0.1875	0.6364	0.3623	0.3538
E1	0.4375	0.3968	0.3846	0.3231	0.3226	0.3438	0.4769	0.3438	0.5224	0.3636	0.3438	0.4848	0.4493	0.2615
E2	0.2381	0.1935	0.1875	0.1250	0.0820	0.1429	0.4375	0.1429	0.3939	0.1692	0.1111	0.5692	0.3235	0.2188
E3	0.1818	0.2308	0.0746	0.0746	0.1250	0.1212	0.4328	0.1818	0.3913	0.0588	0.1212	0.5882	0.2958	0.2836
E4	0.2188	0.2063	0.1692	0.1385	0.0968	0.0938	0.3538	0.1563	0.3433	0.1515	0.1250	0.5758	0.3333	0.2615
E5	0.4154	0.3750	0.3636	0.3030	0.3016	0.2923	0.4242	0.3231	0.4118	0.3433	0.3231	0.4627	0.4000	0.0606
E6	0.2000	0.1875	0.1515	0.1212	0.0794	0.0462	0.3939	0.1385	0.3529	0.1343	0.1077	0.5821	0.3143	0.2727
E9[c]	0.1935	0.1475	0.1746	0.1429	0.0333	0.0968	0.3968	0.0645	0.3538	0.1563	0.0645	0.5625	0.3134	0.2698
E11	0.3750	0.3333	0.3231	0.2615	0.2581	0.2500	0.3846	0.2813	0.4030	0.3030	0.2813	0.4848	0.3623	0.0154
E13	0.1692	0.1563	0.1515	0.1212	0.0794	0.0462	0.4242	0.1385	0.3824	0.1343	0.1077	0.5821	0.3143	0.2727
E14	0.1746	0.1290	0.1563	0.1250	0.0164	0.0794	0.3750	0.0794	0.3333	0.1385	0.0476	0.5385	0.2941	0.2500
E15	0.0769	0.1563	0.1818	0.2121	0.1746	0.2000	0.3939	0.2308	0.4412	0.1940	0.2000	0.6418	0.3714	0.3636

M19	0.1290	0.0000												
M20	0.2258	0.1563	0.0000											
E1	0.2581	0.3438	0.4375	0.0000										
E2	0.1475	0.1429	0.2381	0.3016	0.0000									
E3	0.1875	0.1875	0.1212	0.2121	0.3636	0.1692								
E4	0.1613	0.0938	0.2188	0.3750	0.1746	0.1212	0.0000							
E5	0.3333	0.3231	0.4154	0.2923	0.2813	0.3433	0.3231	0.0000						
E6	0.1429	0.0462	0.2000	0.3538	0.1250	0.1343	0.1077	0.3333	0.0000					
E9[c]	0.1000	0.0968	0.1935	0.3548	0.1148	0.1563	0.0968	0.3333	0.1111	0.0000				
E11	0.3226	0.2813	0.3750	0.2813	0.2381	0.3030	0.2813	0.0462	0.2923	0.2903	0.0000			
E13	0.1429	0.0462	0.1692	0.3538	0.1563	0.1343	0.1077	0.3333	0.0606	0.1111	0.2923	0.0000		
E14	0.0820	0.0794	0.1746	0.3333	0.0645	0.1385	0.1111	0.3125	0.0625	0.0492	0.2698	0.0938	0.0000	
E15	0.2381	0.1692	0.0462	0.4462	0.2188	0.2239	0.2308	0.4242	0.1818	0.2063	0.3846	0.1818	0.1563	0.0000
	M17	M19	M20	E1	E2	E3	E4	E5	E6	E9[c]	E11	E13	E14	E15

The lightly shaded numbers represent the *F* values ≥0.9 (1-the shaded value) that indicate clonally or closely related pairs of isolates that have a one to three fragment difference. The boxed numbers are the 1-*F* values of the most diverse pairs of isolates, M1–M13 and M13–M20, that had an *F* value of 0.36. These pairs shared only 12 of the 54 different fragments found between them.

*The *F* value is 1 minus the value found in this matrix because the RAPDistance Package program used for constructing the matrix has been designed to subtract the *F* value from 1. Therefore, the lowest values in this matrix represent the most closely related pairs of isolates and the highest values represent the most diverse pairs of isolates. The zero values shown in this matrix are the values relating to the comparison of identical pairs that would yield an *F* value of 1 represented by a 0 value in a matrix generated using the RAPDistance Package.

[a]*Sma*I/*Apa*I-M5 included isolates M5, M12, M16, M18 and E12.
[b]*Sma*I/*Apa*I-M8 included isolates M8 and E7.
[c]*Sma*I/*Apa*I-E9 included isolates E9 and E10.

APPENDIX C :

Distribution of 213 Hib isolates among 69 *Sma*I-RFLP types by group and Aboriginality in 7 major genetic distance clusters found using the mathematical model of Nei and Li and the programs, NJTREE and TDRAW

Shared**	Group Name	Repre-senta-tive *Sma*I RFLP type*	Aboriginal	Non-Aboriginal	Total	Page 1 of 2 *refer to Appendices A.1-A.6 for key to code name **a black box indicates a type was found among both Aboriginal and non-Aboriginal isolates
■	G1	NT8	6	2	**8**	NT8, NT11, A5, A7, A16, A19, A37, A40
	F1	M13	0	1	**1**	M13
■	**F2a**	S37	35	2	**37**	S37, A1, A4, A6, A9, A11, A13, A14, A15, A17, A18, A20, A26, A27, A29, A31, A35, A36, A38, A39, A43, A46, A50, WA4, WA6, WA7, WA8, WA9, WA10, WA13, B6, B7, B11, B14, B15, B21, NT4
	F2b	A34	1	0	**1**	A34
	F2c	B5	1	0	**1**	B5
	F3	A23	7	0	**7**	A23, A28, A33, A44, A49, WA2, WA3
	F4	NT12	0	1	**1**	NT12
	E1	M7	0	1	**1**	M7
	E2	V1	0	2	**2**	V1, V9
	E3	M9	0	1	**1**	M9
	E4	T1	0	1	**1**	T1
	A1a	M3	0	4	**4**	M3, M4, WA16, WA25
	A1b	M10	0	2	**2**	M10, S21
	A1c	S10	0	1	**1**	S10
	A2a	E3	0	1	**1**	E3
	A2b	E4	0	1	**1**	E4
	A3a	M6	0	1	**1**	M6
	A3b	M19	0	1	**1**	M19
	A3c	E13	0	1	**1**	E13
	A4a	V8	0	1	**1**	V8
■	A4b	NT2	1	1	**2**	NT2, A47
	A3d	V14	0	1	**1**	V14
	A3e	WA46	0	1	**1**	WA46
	A5a	E6	0	1	**1**	E6
	A5b	S11	0	1	**1**	S11
	A5c	NT10	0	1	**1**	NT10
	A6a	M1	0	1	**1**	M1
	A6b	M20	0	1	**1**	M20
	A6c	E15	0	1	**1**	E15
	A7a	M2	0	2	**2**	M2, S32
	A7b	A42	1	0	**1**	A42
	A7c	A21	1	0	**1**	A21
	A7d	V10	0	1	**1**	V10
	A16	WAA5	1	0	**1**	WA5
	A8a	WA42	0	2	**2**	WA42, NT1

continued on page 2

APPENDIX C :

Distribution of 213 Hib isolates among 69 *Sma*I-RFLP types by group and Aboriginality in 7 major genetic distance clusters found using the mathematical model of Nei and Li and the programs, NJTREE and TDRAW

Shared**	Group Name	Repre-sentative *Sma*I RFLP type*	Aboriginal	Non-Aboriginal	Total	Page 2 of 2 *refer to Appendices A.1-A.6 for key to code name **a black box indicates a type was found among both Aboriginal and non-Aboriginal isolates Shaded isolates are epiglottitis isolates Underlined isolates are those not meningitis or epiglottitis
■	**A8b**	M5	2	53	**55**	M5, M12, M16, M18, E12, V2, V3, V4, V6, V11, V12, V13, V15, V16, V17, V20, S1, S3, S4, S6, S17, S18, S23, S24, S25, S26, S28, S30, S31, S33, S34, S35, S36, WA1, WA14, WA18, WA22, WA23, WA26, WA27, WA28, WA29, WA30, WA32, WA37, WA38, WA39, WA40, WA44, WA45, WA47, NT6, NT9, T2,A22
	A8c	S27	0	1	**1**	S27
	A8d	M11	0	1	**1**	M11
	A9a	M8	0	2	**2**	M8, E7
	A9b	E9	0	3	**3**	E9, E10, WA20
	A10a	M17	0	1	**1**	M17
	A10b	S2	0	1	**1**	S2
	A11a	V5	0	1	**1**	V5
	A11b	S13	0	8	**8**	S13,S20,WA15,WA19,WA31,WA34,WA35, WA43
	A12	A8	1	0	**1**	A8
	A13a	S5	0	1	**1**	S5
	A13b	WA24	0	2	**2**	WA24, WA41
	A14	WA33	0	1	**1**	WA33
	A15a	E2	0	3	**3**	E2, E14, S8
	A15b	V19	0	1	**1**	V19
	B1	B1	1	0	**1**	B1
	B2	E1	0	1	**1**	E1
	B3	V7	0	1	**1**	V7
	B4	S29	0	1	**1**	S29
	B5	V18	0	1	**1**	V18
	B6	M14	0	1	**1**	M14
	B7	S15	0	3	**3**	S15, S16, S22
	C1a	M15	0	3	**3**	M15, S14, S19
	C2	E5	0	1	**1**	E5
	C1b	E11	0	4	**4**	E11, S7, S9, NT7
■	C3	WA17	1	2	**3**	WA17, WA36, WA12
	C4	WA21	0	1	**1**	WA21
	C5	B8	3	0	**3**	B8, B10, B12
	C6	S12	0	1	**1**	S12
	C7a	A45	8	0	**8**	A45, B2, B3, B4, B16, B18, B19, B20
	C7b	B13	1	0	**1**	B13
	D1a	A2	3	0	**3**	A2, A12, A32
	D1b	A41	1	0	**1**	A41
	D2	A3	1	0	**1**	A3
	Totals	**41**	**76**	**137**	**213**	

BIBLIOGRAPHY

Australian Bureau of Statistics. 1987. Census 1986. Canberra, ABS.

Achtman, M. 1994. Clonal spread of serogroup A meningococci: a paradigm for the analysis of microevolution in bacteria. *Molecular Microbiology*, 11: 15-22.

Adams, WG, KA Deaver, SL Cochi, BD Pikaytis, ER Zell, CV Broome and JD Wenger. 1993. Decline of childhood *Haemophilus influenzae* type b (Hib) disease in the Hib vaccine era. *Journal of the American Medical Association*, 269: 221-226.

Adderson, EE, CL Byington, L Spencer, A Kimball, M Hindiyeh, K Carroll, S Mottice, EK Korgenski, JC Christenson and AT Pavia. 2001. Invasive serotype a *Haemophilus influenzae* infections with a virulence genotype resembling *Haemophilus influenzae* type b: emerging pathogen in the vaccine era? *Pediatrics*, 108: 36-42.

Akerley, BJ, EJ Rubin, VL Novick, K Amaya, N Judson and JJ Mekalanos. 2002. A genome-scale analysis for identification of genes required for growth or survival of *Haemophilus influenzae*. *Proceedings of the National Academy of Sciences of the United States of America*, 99: 966-971.

Albritton, WL, S Penner, L Slaney and J Brunton. 1978. Biochemical characteristics of *Haemophilus influenzae* in relationship to source of isolation and antibiotic resistance. *Journal of Clinical Microbiology*, 7: 519-523.

Anderson, EC, NT Begg, SC Crawshaw, RM Hargreaves, AJ Howard and MPE Slack. 1995. Epidemiology of invasive *Haemophilus influenzae* infections in England and Wales in the pre-vaccination era (1990-2). *Epidemiology and Infection*, 115: 89-100.

Arbeit, RD. 1995. Laboratory procedures for the epidemiologic analysis of microorganisms. *In* Manual of Clinical Microbiology. Murray, PR. American Society for Microbiology Press, Washington, DC: 190-208.

Arbeit, RD, R Dunn, J Maslow, R Goldstein and JM Musser. 1990. Resolution of evolutionary divergence and epidemiologic relatedness among *H. influenzae* type b (Hib) by pulsed field gel electrophoresis (PFGE). Program and abstracts of the 30th Interscience Conference on Antimicrobial Agents and Chemotherapy, Atlanta, Georgia, abstract 1144.

Armstrong, JS, AJ Gibbs, R Peakall and G Weiller. 1994. The RAPDistance Package, version 1.04. Australian National University. http://life.anu.edu.au.

Barbour, ML. 1996. Conjugate vaccines and the carriage of *Haemophilus influenzae* type b. *Emerging Infectious Diseases*, 2: 176-182.

Barbour, ML, RT Mayon-White, C Coles, DWM Crook and ER Moxon. 1995. The impact of conjugate vaccines on carriage of *Haemophilus influenzae* type b. *Journal of Infectious Diseases*, 171: 93-98.

Barenkamp, SJ, DM Granoff and RS Munson. 1981. Outer-membrane protein subtypes of *Haemophilus influenzae* type b and spread of disease in day-care centers. *The Journal of Infectious Diseases*, 144: 210-217.

Barenkamp, SJ, DM Granoff and M Pittman. 1983. Outer membrane protein subtypes and biotypes of *Haemophilus influenzae* type b: relation between strains isolated in 1934-1954 and 1977-1980. *The Journal of Infectious Diseases*, 148: 1127.

Barenkamp, SJ, RS Munson and DM Granoff. 1981. Subtyping isolates of *Haemophilus influenzae* type b by outer membrane protein profiles. *Journal of Infectious Diseases*, 143: 668-676.

Barenkamp, SJ, RS Munson and DM Granoff. 1982. Outer membrane protein and biotype analysis of pathogenic nontypeable *Haemophilus influenzae*. *Infection and Immunity*, 36: 535-540.

Bautsch, W. 1992. Bacterial genome mapping by two-dimensional pulsed field gel electrophoresis (2D-PFGE). *In* Methods in Molecular Biology, Volume 12, Pulsed Field Gel Electrophoresis: Protocols, Methods, and Theories. Burmeister, M and L Ulanovsky. The Humana Press, Totowa, NJ: 185-201.

Beni, C, JA Cayla, JM Jansa, R Maldonado, H Panella, L Rajmil and JR Villalbi. 1999. *Haemophilus influenzae* type b invasive disease and meningococcal disease: incidence and main characteristics in Barcelona, Spain, 1994-1995 [Spanish: English abstract]. *Medicina Clinica*, 112: 1-4.

Benson, JA and P Ferrieri. 2001. Rapid pulsed-field gel electrophoresis method for group B streptococcus isolates. *Journal of Clinical Microbiology*, 39: 3006-3008.

Bijlmer, HA, L van Alphen, L Geelen-van den Broek, BM Greenwood, HA Valkenburg and J Dankert. 1992. Molecular epidemiology of *Haemophilus influenzae* type b in the Gambia. *Journal of Clinical Microbiology*, 30: 386-390.

Bio-Rad. 1992. CHEF-DR III Pulsed Field Gel Electrophoresis Systems Instruction Manual and Applications Guide. Bio-Rad Laboratories, Richmond, California.

Booy, R, PT Heath, MPE Slack, N Begg and ER Moxon. 1997. Vaccine failures after primary immunisation with *Haemophilus influenzae* type b conjugate vaccine without booster. *Lancet*, 349: 1197-1202.

Booy, R, S Hodgson, L Carpenter, RT Mayon-White, MPE Slack, JA Macfarlane, EA Haworth, M Kiddle, S Shribman, J St Clair Roberts and ER Moxon. 1994. Efficacy of *Haemophilus influenzae* type b conjugate vaccine PRP-T. *Lancet*, 344: 362-366.

Bower, C, R Condon, J Payne, P Burton, C Watson and B Wild. 1998. Measuring the impact of conjugate vaccines on invasive *Haemophilus influenzae* type b infection in Western Australia. *Australian and New Zealand Journal of Public Health*, 22: 67-72.

Bradshaw, MW, JC Parke, R Schneerson and JB Robbins. 1971. Bacterial antigens cross-reactive with the capsular polysaccharide of *Haemophilus influenzae* type b. *Lancet*, i: 1095-1096.

Breukels, MA, L Spanjaard, LAM Sanders and GT Rijkers. 2001. Immunological characterization of conjugated *Haemophilus influenzae* type b vaccine failure in infants. *Clinical Infectious Diseases*, 32: 1700-1705.

Bruce, KD and JZ Jordens. 1991. Characterization of noncapsulate *Haemophilus influenzae*by whole-cell polypeptide profiles, restriction endonuclease analysis, and rRNA gene restriction patterns. *Journal of Clinical Microbiology*, 29: 291-296.

Bruce, KD and TH Pennington. 1991. Clonal analysis of non-typable *Haemophilus influenzae* by sodium dodecyl sulfate-polyacrylamide gel electrophoresis of whole cell polypeptides. *Journal of Medical Microbiology*, 34: 277-283.

Burroughs, M, S Prasad, C Cabellos, PM Mendelman and E Tuomanen. 1993. The biologic activities of peptidoglycan in experimental *Haemophilus influenzae* meningitis. *Journal of Infectious Diseases*, 167: 464-468.

Busch, U and H Nitschko. 1999. Methods for the differentiation of microorganisms [Review]. *Journal of Chromatography B*, 722: 263-278.

Butler, P and ER Moxon. 1990. A physical map of the genome of *Haemophilus influenzae* type b. *Journal of General Microbiology*, 136: 2333-2342.

Butt, HL, RL Clancy, AW Cripps, K Murree-Allen, NA Saunders, DC Sutherland and MJ Hensley. 1990. Bacterial colonisation of the respiratory tract in chronic bronchitis. *Australian New Zealand Journal of Medicine*, 20: 35-38.

Campagne, G, JP Chippaux, S Djibo, O Issa and A Garba. 1999. Epidemiology and prevention of infants' bacterial meningitis in Niamey, Niger [French: English abstract]. *Bulletin de la Societe de Pathologie Exotique*, 92: 118-122.

Campos, J, M Chanyangam, R deGroot, AL Smith, FC Tenover and R Reig. 1989. Genetic relatedness of antibiotic resistance determinants in multiply resistant *Hemophilus influenzae*. *The Journal of Infectious Diseases*, 160: 810-817.

Campos, JM. 1995. *Haemophilus*. *In* Manual of Clinical Microbiology. Murray, PR. American Society for Microbiology Press, Washington, DC: 556-565.

Campos, JM. 2001. *Haemophilus influenzae*: from the post-vaccination era to antibiotic resistance. *Clinical Microbiology and Infection*, 7: 287-290.

Carlone, GM, ML Thomas, HS Rumschlag and FO Sottnek. 1986. Rapid microprocedure for isolating detergent-insoluable outer membrane proteins from *Haemophilus* species. *Journal of Clinical Microbiology*, 24: 330-332.

Caugant, DA, B-E Kristiansen, LO Froholm, K Bovre and RK Selander. 1988. Clonal diversity of *Neisseria meningitidis* from a population of asymptomatic carriers. *Infection and Immunity*, 56: 2060-2068.

Caugant, DA, LF Mocca, CE Frasch, LO Froholm, WD Sollinger and SR K. 1987a. Genetic structure of *Neisseria meningitidis* populations in relation to serogroup, serotype, and outer membrane protein pattern. *Journal of Bacteriology*, 169: 2781-2792.

Caugant, DA, WD Zollinger, LF Mocca, CE Frasch, TS Whittam, LO Froholm and RK Selander. 1987b. Genetic relationships and clonal population structure of serotype 2 strains of *Neisseria meningitidis*. *Infection and Immunity*, 55: 1503-1513.

Chachaty, E, P Saulnier, A Martin, N Mario and A Andremont. 1994. Comparison of ribotyping, pulsed field gel electrophoresis and random amplified polymorphic DNA for typing *Clostridium difficile* strains. *FEMS Microbiology Letters*, 122: 61-68.

Chang, CC, JR Gilsdorf, VJ DiRita and CF Marrs. 2000. Identification and genetic characterization of *Haemophilus influenzae* genetic island 1. *Infection and Immunity*, 68: 2630-2637.

Chang, N and L Chui. 1998. A standardized protocol for the rapid preparation of bacterial DNA for pulsed field gel electrophoresis. *Diagnostic Microbiology and Infectious Diseases*, 31: 275-279.

Clements, DA and GL Gilbert. 1989. *Haemophilus influenzae* type b infection Victoria and worldwide. *Australian Microbiologist*, 10: 523-526.

Clements, DA, SJ MacInnes and GL Gilbert. 1992. Outer membrane protein subtypes of *Haemophilus influenzae* type b isolates causing invasive disease in Victoria, Australia from 1988 to 1990. *Journal of Clinical Microbiology*, 30: 1879-1881.

Collignon, PJ, JM Bell, SJ MacInnes, GL Gilbert and M Toohey. 1992. A national collaborative study of antimicrobial resistance in *Haemophilus influenzae* in Australian hospitals. *Journal of Antimicrobial Chemotherapy*, 30: 153-163.

Cordova, SP, MT Gillies and MY Beers. 2000. The outbreak that had to happen: Bordetella pertussis in north-west Western Australia in 1999. *Communicable Diseases Intelligence*, 24: 375-379.

Corn, PG, J Anders, AK Takala, H Kayhty and SK Hoiseth. 1993. Genes involved in *Haemophilus influenzae* type b capsule expression are frequently amplified. *Journal of Infectious Diseases*, 167: 356-364.

Crisel, RM, RS Baker and DE Dorman. 1975. Capsular polymer of *Haemophilus influenzae* type b. *Journal of Biological Chemistry*, 250: 4926-4930.

Curran, R, KR Hardie and KJ Towner. 1994. Analysis by pulsed field gel electrophoresis of insertion mutations in the transferrin-binding system of *Haemophilus influenzae* type b. *Journal of Medical Microbiology*, 41: 120-126.

Dagan, R. 1993. Epidemiology of invasive *Haemophilus influenzae* type b (Hib) disease in Israel. *Vaccine*, 11: S43-S45.

Decker, MD and KM Edwards. 1998. *Haemophilus influenzae* type b vaccines: history, choice and comparisons. *Pediatric Infectious Disease Journal*, 17: S113-S116.

Dewhirst, FE, BJ Paster, I Olsen and GJ Raser. 1992. Phylogeny of 54 representative strains of species in the family Pasteurellaceae as determined by comparison of 16S rRNA sequences. *Journal of Bacteriology*, 174: 2002-2013.

Dominguez, A, C Latorre, V Pineda, N Margall, R Bou, D Fontanals, JM Corretger, F Sanchez, T Juncosa, I Santfeliu, J Benet, I Pons, A Martinez, P Ciruela, C Munoz, J Fortea, E Lobera, B Mirelis, J Rello, J Renau, G Prats and L Salleras. 1999.

Haemophilus influenzae type b invasive disease in Catalonia, Spain (1996) [Spanish: English abstract]. *Medicina Clinica*, 112: 5-9.

Dubose, RF, DE Dykhuizen and DL Hartl. 1988. Genetic exchange among natural isolates of bacteria: recombination within the *pho* gene of *Escherichia coli*. *Proceedings of the National Academy of Scientists of the United States of America*,: 7036-7040.

Dykhuizen, DE and L Green. 1991. Recombination in *Escherichia coli* and the definition of biological species. *Journal of Bacteriology*, 173: 7257-7268.

El Saadi, O and AS Cameron. 1993. Epidemiology of *Haemophilus influenzae* type b (Hib) meningitis in South Australia. *Medical Journal of Australia*, 159: 494-495.

El-Adhami, W, JM Kyd, DA Bastin and AW Cripps. 1999. Characterization of the gene encoding a 26-kilodalton protein (OMP26) from nontypeable *Haemophilus influenzae* and immune responses to the recombinant protein. *Infection & Immunity*, 67: 1935-1942.

Ewald, PW. 1994. Evolution of infectious disease. Oxford University Press, New York, New York.

Falkow, S. 1990. The "zen" of bacterial pathogenicity. *In* Molecular Basis of Bacterial Pathogenesis. Iglewski, BH and VL Clark. Academic Press, Inc., San Diego, CA. XI: 3-8.

Falkow, S. 1994. Host-parasite relationships. *In* Sherris: Medical Microbiology. Ryan, KJ. Appleton & Lange, Norwalk, CT: 141-167.

Falla, TJ, DWM Crook, LN Brophy, D Maskell, JS Kroll and ER Moxon. 1994. PCR for capsular typing of *Haemophilus influenzae*. *Journal of Clinical Microbiology*, 32: 2382-2386.

Farber, JM and CJ Addison. 1994. RAPD typing for distinguishing species and strains in the genus *Listeria*. *Journal of Applied Bacteriology*, 77: 242-250.

Farley, MM, DS Stephens, SL Kaplan and EO Mason. 1990. Pilus- and non-pilus mediated interactions of *Haemophilus influenzae* type b with human erythrocytes and human nasopharyngeal mucosa. *Journal of Infectious Diseases*, 161.

Fleischmann, RD, MD Adams, O White, RA Clayton, EF Kirkness, AR Kerlavage, CJ Bult, JF Tomb, BA Dougherty, JM Merrick, K Mckenney, G Sutton, W FitzHugh, C Fields, JD Gocayne, J Scott, R Shirley, LI Liu, A Glodek, JM Kelley, JF Weidman, CA Phillips, T Spriggs, E Hedblom, MD Cotton, TR Utterback, MC Hanna, DT Nguyen, DM Saudek, RC Brandon, LD Fine, JL Fritchman, JL Fuhrmann, NSM Geoghagen, CL Gnehm, LA McDonald, KV Small, CM Fraser, HO Smith and JC Venter. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science*, 269: 496-512.

Flesher, AR and RA Insel. 1978. Characterisation of lipopolysaccharide of *Haemophilus influenzae*. *Journal of Infectious Diseases*, 138: 719-730.

Fogarty, J, AC Moloney and JB Newell. 1995. The epidemiology of *Haemophilus influenzae* type b disease in the Republic of Ireland. *Epidemiology and Infection*, 114: 451-463.

Forbes, KJ, KD Bruce, JZ Jordens, A Ball and TH Pennington. 1991. Rapid methods in bacterial DNA fingerprinting. *Journal of General Microbiology*, 137: 2051-2058.

Fothergill, LD and J Wright. 1933. Influenzal meningitis: the relation of age incidence to the bactericidal power of blood against the causal organism. *Journal of Immunology*, 24: 273-284.

Foxwell, AR, JM Kyd and AW Cripps. 1998. Nontypeable *Haemophilus influenzae*: pathogenesis and prevention [Review]. *Microbiology and Molecular Biology Reviews*, 62: 294-308.

Galil, K, R Singleton, OS Levine, MA Fitzgerald, L Bulkow, M Getty, BA Perkins and A Parkinson. 1999. Reemergence of invasive *Haemophilus influenzae* type b disease in a well-vaccinated population in remote Alaska. *Journal of Infectious Diseases*, 179: 101-106.

Gautom, RK. 1997. Rapid pulsed field gel electrophoresis protocol for typing of *Escherichia coli* O157:H7 and other Gram-negative organisms in 1 day. *Journal of Clinical Microbiology*, 35: 2977-2980.

Gerner-Smidt, P, LM Graves, S Hunter and B Swaminathan. 1998. Computerised analysis of restriction fragment length polymorphism patterns: comparative evaluation of two commercial software packages. *Journal of Clinical Microbiology*, 36: 1318-1323.

Gilbert, GL and DA Clements. 1993. Distribution of *Haemophilus influenzae* type b outer membrane protein subtypes in different ethnic groups in Australia. *Journal of Infectious Diseases*, 168: 255.

Gilbert, GL, DA Clements and SJ Broughton. 1990. *Haemophilus influenzae* type b disease in Victoria, Australia 1985-1987. *Pediatric Infectious Diseases Journal*, 9: 252-257.

Gilbert, GL, P Johnson and DA Clements. 1995. Clinical manifestations and outcome of *Haemophilus influenzae* type b disease. *Journal of Paediatrics and Child Health*, 31: 99-104.

Gilsdorf, JR. 1998. Antigenic diversity and gene polymorphisms in *Haemophilus influenzae* [Review]. *Infection and Immunity*, 66: 5053-5059.

Gilsdorf, JR, HY Chang, KW McCrea and LO Bakaletz. 1992. Comparison of hemagglutinating pili of *Haemophilus influenzae* type b with similar structures of nontypable *H. influenzae*. *Infection and Immunity*, 60: 374-379.

Gilsdorf, JR, KW McCrea and CF Marrs. 1997. Role of pili in *Haemophilus influenzae* adherence and colonization [Review]. *Infection and Immunity*, 65: 2997-3002.

Gomez, E, A Moore, J Sanchez, J Kool, PL Castellanos, JM Feris, M Kolczak and OS Levine. 1998. The epidemiology of *Haemophilus influenzae* type b carriage among infants and young children in Santo Domingo, Dominican Republic. *Pediatric Infectious Disease Journal*, 17: 782-786.

Gomez-de-Leon, P, JI Santos, J Caballero, D Gomez, LE Espinosa, I Moreno, D Pinero and A Cravioto. 2000. Genomic variability of *Haemophilus influenzae* isolated from Mexican children determined by using enterobacterial repetitive intergenic consensus sequences and PCR. *Journal of Clinical Microbiology*, 38: 2504-2511.

Granoff, DM, SJ Barenkamp and RS Munson. 1982. Outer membrane protein subtypes for epidemiological investigation of *Haemophilus influenzae* type b disease. *In Haemophilus influenzae* Epidemiology, Immunology, and Prevention of Disease. Sell, SH and PF Wright. Elsevier Biomedical, New York, New York: 43-54.

Granoff, DM and RS Munson. 1986. Prospects for prevention of *Haemophilus influenzae* type b disease by immunization. *Journal of Infectious Diseases*, 153: 448-461.

Gratten, M and J Montgomery. 1991. The bacteriology of acute pneumonia and meningitis in children in Papua New Guinea: assumptions, facts and technical strategies. *Papua New Guinea Medical Journal*, 34: 185-198.

Gratten, M, F Morey, J Hanna, J Hagget, M Pearson, P Torzillo and J Erlich. 1994. Type, frequency and distribution of *Haemophilus influenzae* in Central Australian Aboriginal children with invasive disease. *Medical Journal of Australia*, 160: 728-729.

Guerina, NG, S Langermann, HW Clegg, HW Kessler, DA Goldmann and JR Gilsdorf. 1982. Adherence of piliated *Haemophilus influenzae* type b to human oropharyngeal cells. *Journal of Infectious Diseases*, 146: 564.

Guthridge, S, P McIntyre, D Isaacs, M Hanlon and M Patel. 2000. Differing serologic responses to an *Haemophilus influenzae* type b polysaccharide-*Neisseria meningitidis* outer membrane protein conjugate (PRP-OMPC) vaccine in Australian Aboriginal and Caucasian infants - implications for disease epidemiology. *Vaccine*, 18: 2584-2591.

Guttman, DS and DE Dykhuizen. 1994. Clonal divergence in *Escherichia coli* as a result of recombination, not mutation. *Science*, 266: 1380-1383.

Hall, LMC. 1994. Are point mutations or DNA rearrangements responsible for the restriction fragment length polymorphisms that are used to type bacteria? *Microbiology*, 140: 197-204.

Hampton, CM, SJ Barenkamp and DM Granoff. 1983. Comparison of outer membrane protein subtypes of *Haemophilus influenzae* type b isolates from healthy children in the general population and diseased patients. *Journal of Clinical Microbiology*,: 596-600.

Hanna, J. 1990. The epidemiology of invasive *Haemophlilus influenzae* infections in children under five years of age in the Northern Territory: a three-year study. *The Medical Journal of Australia*, 152: 234-240.

Hanna, JN. 1992. The epidemiology and prevention of *Haemophilus influenzae* type b in Aboriginal children. *Journal of Paediatric and Child Health*, 28: 352-361.

Hanna, JN and BE Wild. 1991. Bacterial meningitis in children under five years of age in Western Australia. *Medical Journal of Australia*, 155: 160-164.

Hanna, JN, BE Wild and PD Sly. 1992. The epidemiology of acute epiglottitis in children in Western Australia. *Journal of Paediatric and Child Health*, 28: 459-464.

Hansman, D and A Lawrence. 1993. Outer membrane protein and immunoblot analysis of Australian isolates of *Haemophilus influenzae*. *Journal of Medical Microbiology*, 38: 378-383.

Hanson, MS, SE Pelzel, J Latimer, U Mullereberhard and EJ Hansen. 1992. Identification of a genetic locus of *Haemophilus influenzae* type b necessary for the binding and utilization of heme bound to human hemopexin. *Proceedings of the National Academy of Sciences of the United States of America*, 89: 1973-1977.

Hanson, MS, C Slaughter and EJ Hansen. 1992. The *hbpA* gene of *Haemophilus influenzae* type b encodes a heme-binding lipoprotein conserved among heme-dependent *Haemophilus* species. *Infection and Immunity*, 60: 2257-2266.

Hardwick, TH, P Cassiday, RS Weyant, KM Bisgard and GN Sanden. 2002. Changes in predominance and diversity of genomic subtypes of *Bordetella pertussis* isolated in the United States, 1935-1999. *Emerging Infectious Diseases*, 8: 44-49.

Harmsen, D, C Singer, J Rothganger, T Tonjum, GS de Hoog, H Shah, J Albert and M Frosch. 2001. Diagnostics of Neisseriaceae and Moraxellaceae by ribosomal DNA sequencing: ribosomal differentiation of medical microorganisms. *Journal of Clinical Microbiology*, 39: 936-942.

Hartl, DL and DE Dykhuizen. 1984. The population genetics of *Escherichia coli*. *Annual Reviews of Genetics*, 18: 31-68.

Heath, PT. 1998. *Haemophilus influenzae* type b conjugate vaccines: a review of efficacy data. *Pediatric Infectious Disease Journal*, 17: S117-S122.

Heath, TC, MC Hewitt, B Jalaludin, C Roberts, AG Capon, P Jelfs and GL Gilbert. 1997. Invasive *Haemophilus influenzae* type b disease in elderly nursing home residents: two related cases. *Emerging Infectious Diseases*, 3: 179-182.

Herceg, A. 1997. The decline of *Haemophilus influenzae* type b disease in Australia. *Communicable Diseases Intelligence*, 21: 173-175.

High, NJ, ME Deadman and ER Moxon. 1993. The role of a repetitive DNA motif (5'-CAAT-3') in the variable expression of the *Haemophilus influenzae* lipopolysaccharide epitope alpha-Gal(1-4)beta-Gal. *Molecular Microbiology*, 9: 1275-1282.

Hoe, N, K Nakashima, D Grigsby, X Pan, SJ Dou, S Naidich, M Garcia, E Kahn, D Bergmire-Sweat and JM Musser. 1999. Rapid molecular genetic subtyping of serotype M1 group A Streptococcus strains. *Emerging Infectious Diseases*, 5: 254-263.

Hoiseth, SK, PG Corn and J Anders. 1992. Amplification status of capsule genes in *Haemophilus influenzae* type b clinical isolates. *Journal of Infectious Diseases*, 165: S114.

Hoiseth, SK, ER Moxon and RP Silver. 1986. Genes involved in *Haemophilus influenzae* type b capsule expression are part of an 18-kilobase tandem duplication. *Proceedings of the National Academy of Sciences of the United States of America*, 83: 1106-1110.

Holt, JG, NR Kreig, PHA Sneath, JT Staley and ST Williams. 1994. Bergey's Manual of Determinative Bacteriology. Williams & Wilkins.

Hunter, PR and MA Gaston. 1988. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *Journal of Clinical Microbiology*, 26: 2465-2466.

Inglis, B, W El-Adhami and P Stewart. 1993. Methicillin-sensitive and -resistant homologues of *Staphylococcus aureus* occur together among clinical isolates. *Journal of Infectious Diseases*, 167: 323-328.

Inzana, TJ. 1983. Electrophoretic heterogeneity and interstrain variation of the lipopolysaccharide of *H. influenzae*. *Journal of Infectious Diseases*, 148: 492-499.

Janda, JM and SL Abbott. 2002. Bacterial identification for publication: when is enough enough? *Journal of Clinical Microbiology*, 40: 1887-1891.

Jarosik, GP, I Maciver and EJ Hansen. 1995. Utilization of transferrin-bound iron by *Haemophilus influenzae* requires an intact *ton*B gene. *Infection and Immunity*, 63: 710-713.

Jarosik, GP, JD Sanders, LD Cope, U Mullereberhard and EJ Hansen. 1994. A functional *ton*B gene is required for both utilization of heme and virulence expression by *Haemophilus influenzae* type b. *Infection and Immunity*, 62: 2470-2477.

Johnson, P, M Hanlon, D Isaacs and GL Gilbert. 1996. Differing antibody responses to *Haemophilus influenzae* type b after meningitis or epiglottitis. *Epidemiology and Infection*, 116: 21-26.

Johnson, P, F Oppedisano, V Bennett-Wood, GL Gilbert and RM Robins-Browne. 1996. Sporadic invasion of cultured epithelial cells by *Haemophilus influenzae* type b. *Infection and Immunity*, 64: 1051-1053.

Johnson, PDR, SJ MacInnes and GL Gilbert. 1993. Antibodies to *Haemophilus influenzae* type b outer membrane proteins in children with epiglottitis or meningitis and in healthy controls. *Infection and Immunity*, 61: 1531-1537.

Jordens, JZ and MPE Slack. 1995. *Haemophilus influenzae*: then and now. *European Journal of Clinical Microbiology and Infectious Diseases*, 14: 935-948.

Kamiya, H, S Uehara, T Kato, K Shiraki, T Togashi, T Morishima, Y Goto, O Satoh and SM Standaert. 1998. Childhood bacterial meningitis in Japan. *Pediatric Infectious Disease Journal*, 17: S183-S185.

Kauc, L. 1992. Determination of genome size of *Haemophilus influenzae* Sb: analysis of DNA restriction fragments. *Acta Microbiologica Polonica*, 41: 13-24.

Kilian, M. 1976. A taxonomic study of the genus of *Haemophilus*, with the proposal of a new species. *Journal of General Microbiology*, 93: 9-62.

Kilian, M, J Mestecky and RE Schrohenloher. 1979. Pathogenic species of the genus *Haemophilus* and *Streptococcus pneumoniae* produce immunoglobulin A1 protease. *Infection and Immunity*, 26: 143-149.

Killian, M and EL Biberstein. 1984. Genus II. Haemophilus. *In* Bergey's Manual of Systematic Bacteriology. Kreig, NR and JG Holt. Williams & Wilkins, Baltimore. 1: 558-569.

Kimura, A and EJ Hansen. 1986. Antigenic and phenotypic variations of *Haemophilus influenzae* type b lipopolysaccharide and their relationship to virulence. *Infection and Immunity*, 51: 69-70.

Kolbert, CP and DH Persing. 1999. Ribosomal DNA sequencing as a tool for identification of bacterial pathogens [Review]. *Current Opinion in Microbiology.*, 2: 299-305.

Koneman, EW, SD Allen, VR Dowell, WM Janda, HM Sommers and WC Winn. 1988. *Haemophilus*. *In* Color Atlas and Textbook of Diagnostic Microbiology. J B Lippincot Company, Philadelphia, Pennsylvania: 269-290.

Kristensen, K. 1999. Thesis - Special Report. *Haemophilus influenzae* type b: epidemiology of invasive diseases, antimicrobial resistance of clinical isolates, and response to a conjugate vaccine in selected risk groups. Copenhagen University. Copenhagen, Denmark.

Kroll, J and E Moxon. 1988. Capsulation and gene copy number at the *cap* locus of *Haemophilus influenzae* type b. *Journal of Bacteriology*, 170: 859-864.

Kroll, JS. 1992. The genetics of encapsulation in *Haemophilus influenzae*. *Journal of Infectious Diseases*, 165: S93-S96.

Kroll, JS, S Ely and ER Moxon. 1991. Capsular typing of *Haemophilus influenzae* with a DNA probe. *Molecular and Cellular Probes*, 5: 375-379.

Kuklinska, D and M Kilian. 1984. Relative proportions of *Haemophilus* species in the throat of healthy children and adults. *European Journal of Clinical Microbiology*, 3: 249-252.

Kyd, JM and AW Cripps. 1998. Potential of a novel protein, OMP26, from nontypeable *Haemophilus influenzae* to enhance pulmonary clearance in a rat model. *Infection & Immunity*, 66: 2272-2278.

Lai, E. 1991. Pulsed field gel electrophoresis: an overview. *The New England Biolabs Transcript*, 3: 1-2.

Langford, PR, BJ Sheehan, T Shaikh and JS Kroll. 2002. Active copper- and zinc-containing superoxide dismutase in the cryptic genospecies of *Haemophilus* causing urogenital and neonatal infections discriminates them from *Haemophilus influenzae* sensu stricto. *Journal of Clinical Microbiology*, 40: 268-270.

Lau, YL, L Low, R Yung, KW Ng, CW Leung, WH Lee, A Ho and SJ Oppenheimer. 1995. Invasive *Haemophilus influenzae* type b infections in children hospitalized in Hong Kong, 1986-1990. *Acta Paediatrica*, 84: 173-176.

Lau, YL, R Yung, L Low, R Sung, CW Leung and WH Lee. 1998. *Haemophilus influenzae* type b infections in Hong Kong. *Pediatric Infectious Disease Journal*, 17: S165-S169.

Leach, AJ, JB Boswell, V Asche, TG Nienhuys and JD Mathews. 1994. Bacterial colonization of the nasopharynx predicts very early onset and persistence of otitis media in Australian Aboriginal infants. *Pediatric Infectious Disease Journal*, 13: 983-989.

Leaves, NI, M Barbour, JZ Jordens, TEA Peto and DWM Crook. 1994. Invasive-disease-causing and nasopharyngeal isolates of *Haemophilus influenzae* type b are different. *In* Haemophilus, Acintobacillus, and Pasteurella. Donachie, W, FA Lainson and JC Hodgson. Plenum Press, New York: 210-211.

Leaves, NI and JZ Jordens. 1994. Development of a ribotyping scheme for *Haemophilus influenzae* type b. *European Journal of Clinical Microbiology and Infectious Diseases*, 13: 1038-1045.

Lee, BC. 1992. Isolation of an outer membrane hemin-binding protein of *Haemophilus influenzae* type b. *Infection and Immunity*, 60: 810-816.

Lee, HJ. 1998. Epidemiology of systemic *Haemophilus influenzae* disease in Korean children. *Pediatric Infectious Disease Journal*, 17: S185-S189.

Lee, JJ and HO Smith. 1988. Sizing of the *Haemophilus influenzae* Rd genome by pulsed field agarose gel electrophoresis. *Journal of Bacteriology*, 170: 4402-4405.

Lee, YS, G Kumarasinghe, C Chow, E Khor and BW Lee. 2000. Invasive *Haemophilus influenzae* type b infections in Singapore children: a hospital-based study. *Journal of Paediatrics and Child Health*, 36: 125-127.

Levine, OS, G Liu, RL Garman, SF Dowell, SG Yu and YH Yang. 2000. *Haemophilus influenzae* type b and *Streptococcus pneumoniae* as causes of pneumonia among children in Beijing, China. *Emerging Infectious Diseases*, 6: 165-170.

Levy, J, G Verhaegen, P Demol, M Couturier, D Dekegel and JP Butzler. 1993. Molecular characterization of resistance plasmids in epidemiologically unrelated strains of multiresistant *Haemophilus influenzae*. *Journal of Infectious Diseases*, 168: 177-187.

LiPuma, JJ, H Richman and TL Stull. 1990. Haemocin, the bacteriocin produced by *Haemophilus influenzae* type b: species distribution and role in colonization. *Infection and Immunity*, 58: 1600-1605.

Loeb, M, AL Zachary and DH Smith. 1981. Isolation and partial characterization of outer and inner membranes from encapsulated *Haemophilus influenzae* type b. *Journal of Bacteriology*, 145: 596-604.

Loeb, MR and DH Smith. 1980. Outer membrane protein composition in disease isolates of *Haemophilus influenzae*: pathogenic and epidemiological implications. *Infection and Immunity*, 30: 709-717.

Losonsky, G, M Santosahm, V Sehgal, A Zwahlen and E Moxon. 1984. *Haemophilus influenzae* disease in the White Mountain Apaches: molecular epidemiology of a high risk population. *Pediatric Infectious Diseases*, 3: 539-547.

Madore, DV. 1996. Impact of immunization on *Haemophilus influenzae* type b disease [Review]. *Infectious Agents and Disease*, 5: 8-20.

Maiden, MCJ, JA Bygraves, E Feil, G Morelli, JE Russell, R Urwin, Q Zhang, JJ Zhou, K Zurth, DA Caugant, IM Feavers, M Achtman and BG Spratt. 1998. Multilocus sequence typing - a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proceedings of the National Academy of Sciences of the United States of America*, 95: 3140-3145.

Maniatis, T, J Sambrook and EF Fritsch. 1989. Molecular Cloning: A Laboratory Manual, 2nd Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Marrs, CF, GP Krasan, KW McCrea, DL Clemans and JR Gilsdorf. 2001. *Haemophilus influenzae* - human specific bacteria [Review]. *Frontiers in Bioscience*, 6: E41-E60.

Martin, K, G Morlin, A Smith, A Nordyke, A Eisenstark and M Golomb. 1998. The tryptophanase gene cluster of *Haemophilus influenzae* type b: evidence for horizontal gene transfer. *Journal of Bacteriology*, 180: 107-118.

Maskell, DJ, MJ Szabo, ME Deadman and ER Moxon. 1992. The *gal* locus from *Haemophilus influenzae*: cloning, sequencing and the use of *gal* mutants to study lipopolysaccharide. *Molecular Microbiology*, 6: 3051-3063.

Maslow, JN, AM Slutsky and RD Arbeit. 1993. Application of pulsed field gel electrophoresis to molecular epidemiology. *In* Diagnostic Molecular Microbiology: principles and applications. Persing, DH, TF Smith, FC Tenover and TJ White. American Society for Microbiology Press, Washington, DC: 563-572.

Matushek, MG, MJM Bonten and MK Hayden. 1996. Rapid preparation of bacterial DNA for pulsed field gel electrophoresis. *Journal of Clinical Microbiology*, 34: 2598-2600.

Maynard Smith, J. 1990. The Evolution of Prokaryotes: Does Sex Matter? *Annual Reviews Ecological Systems*, 21: 1-12.

Maynard Smith, J. 1993a. The role of sex in bacterial evolution. *Journal of Heredity*, 84: 326-327.

Maynard Smith, J, CG Dowson and BG Spratt. 1991. Localized sex in bacteria. *Nature*, 349: 29-31.

Maynard Smith, J, NH Smith, M O'Rourke and BG Spratt. 1993b. How clonal are bacteria? *Proceedings of the National Academy of Science USA*, 90: 4384-4388.

McGee, L, L McDougal, J Zhou, BG Spratt, FC Tenover, R George, R Hakenbeck, W Hryniewicz, JC Lefevre, A Tomasz and KP Klugman. 2001. Nomenclature of major antimicrobial-resistant clones of *Streptococcus pneumoniae* defined by the pneumococcal molecular epidemiology network. *Journal of Clinical Microbiology*, 39: 2565-2571.

McGregor, AR, JM Bell, IM Abdool and PJ Collignon. 1992. Invasive *Haemophilus influenzae* infection in the Australian Capital Territory region. *Medical Journal of Australia*, 156: 569-572.

McIntyre, PB, T Chey and WT Smith. 1995. The impact of vaccination against invasive *Haemophilus influenzae* type b disease in the Sydney region. *Medical Journal of Australia*, 162: 245-248.

McIntyre, PB, SR Leeder and LM Irwig. 1991. Invasive *Haemophilus influenzae* type b in Sydney children 1985-1987: a population-based study. *Medical Journal of Australia*, 154: 832-837.

Melville, CAS, IF Laurenson and JA Augrove. 1994. Invasive disease due to *Haemophilus influenzae* infection. *British Medical Journal*, 309: 58.

Mencarelli, M, C Marsili, A Zanchi, C Pantini and C Cellesi. 1993. Genomic DNA fingerprints and phenotypic characteristics of serotype b *Haemophilus influenzae* isolates from Italy. *European Journal of Epidemiology*, 9: 353-360.

Mendelman, PM, VP Syriopoulou, SL Gandy, JI Ward and AL Smith. 1985. Molecular epidemiology of plasmid-mediated ampicillin resistance in *Haemophilus influenzae* type b isolates from Alaska. *Journal of Infectious Diseases*, 151: 1061-1072.

Michael, A, B Combs, M Gratten, J Montgomery, T Lupiwa, J Mgone and D Lehmann. 1997. Emergence of multiresistant *Haemophilus influenzae* type b in Papua New Guinea. International Conference on Acute Respiratory Infections, Canberra, ACT.

Michaels, RH and CW Norden. 1977. Pharyngeal colonization with *Haemophilus influenzae* type b: a longitudinal study of families with a child with meningitis or epiglottitis due to *H. influenzae* type b. *Journal of Infectious Disease*, 136: 222-228.

Michaels, RH, CS Poziviak, FE Stonebraker and CW Norden. 1976. Factors affecting pharyngeal *Haemophilus influenzae* type b colonization rates in children. *Journal of Clinical Microbiology*, 4: 413-417.

Milkman, R. 1973. Electrophoretic variation in *Escherichia coli* from natural sources. *Science*, 182: 1024-1026.

Milkman, R and A Stolzfus. 1988. Molecular evolution of the *Escherichia coli* chromosome II. Clonal segments. *Genetics*, 120: 359-366.

Mitchell, D and V Krause. 1995. Evaluation of the Northern Territory's *Haemophilus influenzae* type b (Hib) vaccine promotion campaign. *Communicable Diseases Intelligence*, 19: 93-96.

Mitsuda, T, H Kuroki, N Ishikawa, T Imagawa, S Ito, T Miyamae, M Mori, S Uehara and S Yokota. 1999. Molecular epidemiological study of *Haemophilus influenzae* serotype b strains obtained from children with meningitis in Japan. *Journal of Clinical Microbiology*, 37: 2548-2552.

Miyazaki, S, T Nunoya, T Matsumoto, N Furuya, K Tateda and K Yamaguchi. 2001. Lipoolygosaccharide indirectly enhances inflammatory lesions in lungs as a primary infection site by non-encapsulated and type b *Haemophilus influenzae* through production of cytokines. *Cytokine*, 15: 171-174.

Mooi, FR, H van Oirschot, K Heulvelman, HJG van der Heide, W Gaastra and RJL Willems. 1998. Polymorphism in the *Bordetella pertussis* virulence factors P.69/pertactin and pertussis toxin in The Netherlands: temporal trends and evidence for vaccine driven evolution. *Infection and Immunity*, 66: 670-675.

Moor, PE and R Kelly. 1995. The identification of invasive *Haemophilus influenzae* using molecular techniques. Australian Society for Microbiology and the New Zealand Microbiological Society Combined Scientific Meeting and Exhibition, Canberra.

Morgenroth, A and JP Duguid. 1968. Demonstration of different mutational sites controlling rhamnose fermentation in FIRN and non-FIRN rha– strains of *Salmonella typhimurium*: an essay in bacterial archaeology. *Genetical Research (Cambridge)*, 11.

Moxon, E, A Smith, D Averill and D Smith. 1974. *Haemophilus influenzae* meningitis in infant rats after intranasal inoculation. *Journal of Infectious Diseases*, 129: 154-162.

Moxon, ER. 1986. The carrier state: *Haemophilus influenzae*. *Journal of Antimicrobial Chemotherapy*, 18: 17-24.

Moxon, ER and JS Kroll. 1990. The role of bacterial polysaccharide capsules as virulence factors. *Current Topics in Microbiology and Immunology*, 150: 65-85.

Moxon, ER and PA Murphy. 1978. *Haemophilus influenzae* bacteremia and meningitis resulting from survival of a single organism. *Proceedings of the National Academy of Scientists of the United States of America*, 75: 1534 -1536.

Moxon, ER and PT Ostrow. 1977. *Haemophilus influenzae* meningitis in rats: the role of bacteremia in the pathogenesis of age-dependent inflammatory responses in cerebrospinal fluid. *Journal of Infectious Diseases*, 135: 303-307.

Moxon, ER and KA Vaughn. 1981. The type b capsular polysaccharide as a virulence determinant of *Haemophilus influenzae*: studies using clinical isolates and laboratory transformants. *Journal of Infectious Diseases*, 143: 517-524.

Muhlemann, K, M Balz, S Aebi and K Schopfer. 1996. Molecular characteristics of *Haemophilus influenzae* causing invasive disease during the period of vaccination in Switzerland: analysis of strains isolated between 1986 and 1993. *Journal of Clinical Microbiology*, 34: 560-563.

Mulholland, K, S Hilton, R Adegbola, S Usen, A Oparaugo, C Omosigho, M Weber, A Palmer, G Schneider, K Jobe, G Lahai, S Jaffar, O Secka, K Lin, C Ethevenaux and B Greenwood. 1997. Randomised trial of *Haemophilus influenzae* type b tetanus protein conjugate for prevention of pneumonia and meningitis in Gambian infants. *Lancet*, 349: 1191-1197.

Mulvey, MR, L Chui, J Ismail, L Louie, C Murphy, N Chang and M Alfa. 2001. Development of a Canadian standardized protocol for subtyping methicillin-resistant *Staphylococcus aureus* using pulsed field electrophoresis. *Journal of Clinical Microbiology*, 39: 3481-3485.

Munson, E, M Pfaller, F Koontz and G Doern. 2002. Comparison of porphyrin-based, growth factor-based, and biochemical-based testing methods for identification of *Haemophilus influenzae*. *European Journal of Clinical Microbiology and Infectious Diseases*, 21: 196-203.

Munson, RJ, C Bailey and S Grass. 1989. Diversity of the outer membrane protein P2 gene from major clones of *Haemophilus influenzae* type b. *Molecular Microbiology*, 3: 1797-1803.

Murley, YM, TD Edlind, PA Plett and JJ Lipuma. 1998. Cloning of the haemocin locus of *Haemophilus influenzae* type b and assessment of the role of haemocin in virulence. *Microbiology*, 144: 2531-2538.

Murley, YM, TD Edlind, JM Pozsgay and JJ Lipuma. 1997. Cloning and characterization of the haemocin immunity gene of *Haemophilus influenzae*. *Journal of Bacteriology*, 179: 1684-1689.

Murphy, TF and MA Apicella. 1987. Nontypable *Haemophilus influenzae*: a review of clinical aspects, surface antigens, and the human immune response to infection. *Reviews of Infectious Diseases*, 9: 1-15.

Murphy, TV, MT Osterholm, LM Pierson, KE White, JA Breedlove, GB Seibert, JN Kuritsky and DM Granoff. 1987. Prospective surveillance of *Haemophilus influenzae* type b disease in Dallas County, Texas and Minnesota. *Pediatrics*, 79: 173-180.

Musser, J, JS Kroll, DM Granoff, ER Moxon, BR Brodeur, J Campos, H Dabernat, W Frederiksen, J Hamel, G Hammond, EA Hoiby, KR Jonsdottir, M Kabeer, I Kallings, WN Khan, M Kilian, K Knowles, HJ Koornhof, B Law, KI Li, J Montgomery, PE Pattison, J-C Piffaretti, AK Takala, ML Thong, RA Wall, JI Ward and RK Selandar. 1990. Global genetic structure and molecular epidemiology of encapsulated *Haemophilus influenzae*. *Reviews of Infectious Diseases*, 12: 75-111.

Musser, JM. 1996. Molecular population genetic analysis of emerged bacterial pathogens: selected insights. *Emerging Infectious Diseases*, 2: 1-17.

Musser, JM, DM Granoff, PE Pattison and RK Selander. 1985. A population genetic framework for the study of invasive diseases caused by serotype b strains of *Haemophilus influenzae*. *Proceedings of the National Academy of Sciences of the United States of America*, 82: 5078-5082.

Musser, JM, V Kapur, J Szeto, X Pan, DS Swanson and DR Martin. 1995. Genetic diversity and relationships among *Streptococcus pyogenes* strains expressing serotype M1 protein - recent intercontinental spread of a subclone causing episodes of invasive disease. *Infection & Immunity*, 63: 994-1003.

Musser, JM, JS Kroll, ER Moxon and RK Selander. 1988. Clonal population structure of encapsulated *Haemophilus influenzae*. *Infection and Immunity*, 56: 1837-1845.

Musser, JM, JS Kroll, ER Moxon and RK Selander. 1988. Evolutionary genetics of the encapsulated strains of *Haemophilus influenzae*. *Proceedings of the National Academy of Science of the United States of America*, 85: 7758-7762.

Nei, M and WH Li. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America*, 76: 5269-5273.

Neuzil, KM. 2000. Eradication of polio, measles, and Hib. Third Annual Conference on Vaccine Research.

Noel, GJ, SK Hoiseth and PJ Edelson. 1992. Type b capsule inhibits ingestion of *Haemophilus influenzae* by murine macrophages: studies with isogenic encapsulated and nonencapsulated strains. *Journal of Infectious Diseases*, 166: 178-182.

Olive, DM and P Bean. 1999. Principles and applications of methods for DNA-based typing of microbial organisms [Review]. *Journal of Clinical Microbiology.*, 37: 1661-1669.

Olyhoek, T, BA Crowe and M Achtman. 1987. Clonal population structure of *Neisseria meningitidis* serogroup A isolated from epidemics and pandemics between 1915 and 1983. *Reviews of Infectious Diseases*, 9: 655-691.

Orskov, F and I Orskov. 1983. Summary of a workshop on the clone concept in the epidemiology, taxonomy, and evolution of the Enterobacteriaceae and other bacteria. *The Journal of Infectious Diseases*, 148: 346-357.

Orskov, F, V Orskov, DJ Evans, RB Sack, DA Sack and T Waldstrom. 1976. Special *Escherichia coli* serotypes among enterotoxigenic strains from diarrhoea in adults and children. *Medical Microbiology and Immunology*, 162: 73-80.

Paterson, AJ, KF MacSween and TH Pennington. 1987. *Haemophilus influenzae* subtyping by SDS-PAGE of whole-cell polypeptides. *Epidemiology and Infection*, 99: 179-189.

Peltola, H. 1993. *H. influenzae* in the post-vaccination era. *Lancet*, 341: 864-865.

Peltola, H. 2000. Worldwide *Haemophilus influenza*e type b disease at the beginning of the 21st century: global analysis of the disease burden 25 years after the use of the polysaccharide vaccine and a decade after the advent of conjugates. *Clinical Microbiology Reviews*, 13: 302-317.

Peltola, H, H Kayhty, A Sivonen and PH Makela. 1977. *Haemophilus influenzae* type b capsular polysaccharide vaccine in children: a doubleblind field study of 100,000 vaccinees 3 months to 5 years of age in Finland. *Pediatrics*, 60: 730-737.

Peltola, H, T Kilpi and M Anttila. 1992. Rapid disappearance of *Haemophilus influenzae* type b meningitis after routine childhood immunization with conjugate vaccines. *Lancet*, 340: 592-594.

Pennington, TH. 1994. Molecular systematics and traditional medical microbiologists - problems and solutions. *Journal of Medical Microbiology*, 41: 371-373.

Pennington, TH and EM Freebairn. 1989. Subtyping of *Haemophilus influenzae* type b strains from Europe and North America by SDS-PAGE of whole-cell polypeptides: the geographical distribution of subtypes. *Epidemiology and Infection*, 102: 11-19.

Pettigrew, MM, B Foxman, Z Ecevit, CF Marrs and J Gilsdorf. 2002. Use of pulsed-field gel electrophoresis, enterobacterial repetitive intergenic consensus typing, and automated ribotyping to assess genomic variability among strains of nontypeable *Haemophilus influenzae*. *Journal of Clinical Microbiology*, 40: 660-662.

Pfeiffer, R. 1893. Die aetiologi der influenza. *Zeitschrift fur Hygiene und Infektionskrank-heiten*, 13: 357-386.

Pichichero, ME, J Loeb, P Anderson and SH Smith. 1982. Do pili play a role in pathogenicity of *Haemophilus influenzae* type b? *Lancet*, ii: 960-962.

Pittman, M. 1931. Variation and type specificity in the bacterial species *Haemophilus influenzae*. *Journal of Experimental Medicine*, 53: 471-492.

Plaut, AG. 1983. The IgA1 proteases of pathogenic bacteria. *Annual Reviews of Microbiology*, 37: 603-622.

Popovic, T, S Schmink, NA Rosenstein, GW Ajello, MW Reeves, B Plikaytis, SB Hunter, EM Ribot, D Boxrud, ML Tondella, C Kim, C Noble, E Mothershed, J Besser and BA Perkins. 2001. Evaluation of pulsed field gel electrophoresis in epidemiological investigations of meningococcal disease outbreaks caused by *Neisseria meningitidis* serogroup C. *Journal of Clinical Microbiology*, 39: 75-85.

Porras, O, DA Caugant, B Gray, T Lagergard, BR Levin and C Svanborg-Eden. 1986. Difference in structure between type b and nontypable *Haemophilus influenzae* populations. *Infection and Immunity*, 53: 79-89.

Preston, A and MA Apicella. 1999. The genetics of capsule and lipooligosaccharide biosynthesis in *Haemophilus influenzae*. *In* Genetics of Bacterial Polysaccharides. CRC Press: 91-109.

Quentin, R, C Martin, JM Musser, N Pasquierpicard and A Goudeau. 1993. Genetic characterization of a cryptic genospecies of *Haemophilus* causing urogenital and neonatal infections. *Journal of Clinical Microbiology*, 31: 1111-1116.

Reich, KA, L Chovan and P Hessler. 1999. Genome scanning in *Haemophilus influenzae* for identification of essential genes. *Journal of Bacteriology*, 181: 4961-4968.

Reinert, P, A Liwartowski, H Dabernat, C Guyot, J Boucher and C Carrere. 1993. Epidemiology of *Haemophilus influenzae* type b disease in France. *Vaccine*, 11: S38-S42.

Renders, N, U Romling, H Verbrugh and A van Belkum. 1996. Comparative typing of *Pseudomonas aeruginosa* by random amplification of polymorphic DNA or pulsed field gel electrophoresis of DNA macrorestriction fragments. *Journal of Clinical Microbiology*, 34: 3190-3195.

Rietschel, ET and H Brade. 1992. Bacterial endotoxins. *Scientific American*, 267: 26-33.

Roord, JJ, M Apicella and WM Scheld. 1994. The induction of meningeal inflammation and blood-brain barrier permeability by *Haemophilus influenzae* type b peptidoglycan. *Journal of Infectious Diseases*, 170: 254-256.

Rubin, I and ER Moxon. 1983. Pathogenesis of bloodstream invasion with *Haemophilus influenzae* type b. *Infection and Immunity*, 41: 280-284.

Ryan, KJ and S Falkow. 1994. Haemophilus and Bordetella. *In* Sherris Medical Microbiology. Ryan, KJ. Appleton & Lange, Norwalk, Connecticut: 362-363.

Saito, M, A Umeda and S Yoshida. 1999. Subtyping of *Haemophilus influenzae* strains by pulsed field gel electrophoresis. *Journal of Clinical Microbiology*, 37: 2142-2147.

Saitou, N and M Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4: 406-425.

Sambrook, J and DW Russell. 2001. Molecular Cloning: A Laboratory Manual, 3rd Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sandery, M, J Coble and S McKersie-Donnolley. 1994. Random amplified polymorphic DNA (RAPD) profiling of *Legionella pneumophila*. *Letters in Applied Microbiology*, 19: 184-187.

Scheifele, D. 2001. Hib vaccines: lessons learned. *In* American Medical Association Continuing Education Program: Impact of Conjugate Vaccines on Paediatric Disease. Oxford Institute for Continuing Education, http://www.medscape.com/viewarticle /413130_17.

Scheifele, DW, TP Jadavji, BJ Law, R Gold, NE Macdonald, MH Lebel, EL Mills, P Dery, SA Halperin, RF Morris, V Marchessault and PJ Duclos. 1996. Recent trends in pediatric *Haemophilus influenzae* type b infections in Canada. *Canadian Medical Association Journal*, 154: 1041-1047.

Schwartz, DC and CR Cantor. 1984. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell*, 37: 67-75.

Scott, J. 1993. Review of notifications of invasive *Haemophilus influenzae* type b infections in Queensland, 1990-1993. *Communicable Diseases Intelligence*, 17: 403-406.

Selander, RK, DA Caugant, H Ochman, JM Musser, MN Gilmour and TS Whittam. 1986. Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. *Applied and Environmental Microbiology*, 51: 873-884.

Selander, RK and BR Levin. 1980. Genetic diversity and structure in *Escherichia coli* populations. *Science*, 210: 545-547.

Selander, RK, JM Musser, DA Caugant, MN Gilmour and TS Whittam. 1987. Population genetics of pathogenic bacteria. *Microbial Pathogenesis*, 3: 1-7.

Shapiro, ED and JI Ward. 1991. The epidemiology and prevention of disease caused by *Haemophilus influenzae* type b. *Epidemiologic Reviews*, 13: 113-142.

Slack, MPE. 1995. Invasive *Haemophilus influenzae* disease: the impact of Hib immunisation. *Journal of Medical Microbiology*, 42: 75-77.

Smith, HO and KW Wilcox. 1970. A restriction enzyme from *Haemophilus influenzae*. *Journal of Molecular Biology*, 51: 379-391.

Smith-Vaughan, HC, KS Sriprakash, AJ Leach, JD Mathews and DJ Kemp. 1998. Low genetic diversity of *Haemophilus influenzae* type b compared to nonencapsulated *H. influenzae* in a population in which *H. influenzae* is highly endemic. *Infection and Immunity*, 66: 3403-3409.

Smith-Vaughan, HC, KS Sriprakash, JD Mathews and DJ Kemp. 1995. Long PCR-ribotyping of nontypeable *Haemophilus influenzae*. *Journal of Clinical Microbiology*, 33: 1192-1195.

Spratt, BG, WP Hanage and EJ Feil. 2001. The relative contributions of recombination and point mutation to the diversification of bacterial clones [Review]. *Current Opinion in Microbiology*, 4: 602-606.

Spratt, BG and MCJ Maiden. 1999. Bacterial population genetics, evolution and epidemiology. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences.*, 354: 701-710.

St. Geme, JW. 1994. Early events in the pathogenesis of *Haemophilus influenzae* disease. *In* Molecular Genetics of Bacterial Pathogenesis. Miller, VL, JB Kaper, DA Portnoy and RR Isberg. American Society for Microbiology Press, Washington, DC: 157-172.

St. Geme, JW and S Falkow. 1991. Loss of capsule expression by *Haemophilus influenzae* type b results in enhanced adherence to and invasion of human cells. *Infection and Immunity*, 59: 1325-1333.

Staley, JT and NR Krieg. 1984. Classification of prokaryote organisms: an overview. *In* Bergey's Manual of Systematic Bacteriology. Krieg, NR and JG Holt. Williams & Wilkins, Baltimore.

Stewart, PR, W El-Adhami and B Inglis. 1994. Pulsed field gel electrophoresis: applications to bacteria. *In* Methods in Gene Technology. JAI Press. 2: 177-205.

Strausbaugh, LJ. 1997. *Haemophilus influenzae* infections in adults. *Postgraduate Medicine*, 101: 191-200.

Struelens, MJ. 1998. Molecular epidemiologic typing systems of bacterial pathogens: current issues and perspectives. *Memorias do Instituto Oswaldo Cruz*, 93: 581-585.

Stull, TL, JJ LiPuma and TD Edind. 1988. A broad-spectrum probe for molecular epidemiology of bacteria: ribsomal RNA. *The Journal of Infectious Diseases*, 157: 280-286.

Sutton, A, R Schneerson, S Kendall-Morris and JB Robbins. 1982. Differential complement resistance mediates virulence of *Haemophilus influenzae* type b. *Infection and Immunity*, 35: 95-104.

Swaminathan, B and GM Matar. 1993. Molecular typing methods. *In* Diagnostic Molecular Biology Principles and Applications. Persing, DH, TF Smith, FC Tenover and TJ White. American Society for Microbiology Press, Washington, DC: 26-50.

Swanson, DS, X Pan, MW Kline, RE McKinney, R Yogev, LL Lewis, MT Brady, GD McSherry, WM Dankner and JM Musser. 1998. Genetic diversity among *Mycobacterium avium* complex strains recovered from children with and without human immunodeficiency virus infection. *Journal of Infectious Diseases*, 178: 776-782.

Takala, AK, J Eskola, P Bol, L van Alphen, J Palmgren and PH Makela. 1987. *Haemophilus influenzae* type b strains of outer membrane subtypes 1 and 1c cause different types of invasive disease. *Lancet*, 2: 647-650.

Takala, AK, J Eskola, H Peltola and H Makela. 1989. Epidemiology of invasive *Haemophilus influenzae* type b disease among children in Finland before vaccination with *Haemophilus influenzae* type b vaccine. *Pediatric Infectious Disease Journal*, 8: 297-302.

Takala, AK, L van Alphen, JM Musser, L Geelen, RK Selander, J Eskola and PH Makela. 1989. Bacteriologic epidemiology of *Haemophilus influenzae* type b strains causing invasive infection in Finland. *Journal of Infectious Diseases*, 148: 75-81.

Tarasi, A, F D'Ambrosia, G Perrone and A Pantosti. 1998. Susceptibility and genetic relatedness of invasive *Haemophilus influenzae* type b in Italy. *Microbial Drug Resistance*, 4: 301-306.

Taylor, JW, DM Geiser, A Burt and V Koufopanooou. 1999. The evolutionary biology and population genetics underlying fungal strain typing. *Clinical Microbiology Reviews*, 12: 126-146.

Tenover, FC, RD Arbeit, RV Goering, PA Mickelsen, BE Murray, DH Persing and B Swaminathan. 1995. Interpreting chromosomal DNA restriction patterns produced by pulsed field gel electrophoresis: criteria for bacterial strain typing. *Journal of Clinical Microbiology*, 33: 2233-2239.

Thal, LA, J Silverman, S Donabedian and MJ Zervos. 1997. The effect of Tn*916* insertions on contour-clamped homogeneous electrophoresis patterns of *Enterococcus faecalis*. *Journal of Clinical Microbiology*, 35: 969-972.

Tsolia, MN, RN Theodoridou, GJ Mostrou, Paraskaki, II, AM Pangali, AS Yelesme, PK Kalambalikis, AE Gaviotaki, DJ Zoumboulakis and CA Sinaniotis. 1998. Epidemiology of invasive *Haemophilus influenzae* type b infections among children in Greece before the introduction of immunization. *Scandinavian Journal of Infectious Diseases*, 30: 165-168.

Turk, DC. 1981. *Haemophilus influenzae*. *In* Public Health Laboratory Service Monograph Series. Public Health Laboratory Service Publications, HMSO, London. Number 17.

Turk, DC. 1982a. Clinical importance of *Haemophilus influenzae*. *In Haemophilus influenzae* Epidemiology, Immunology, and Prevention of Disease. Sell, SH and PF Wright. Elsevier Biomedical, New York, New York: 3-9.

Turk, DC. 1982b. *Haemophilus influenzae*. *In* Public Health Laboratory Service Monograph Series. Public Health Laboratory Service Publications, HMSO, London. Number 17.

Turk, DC. 1984. The pathogenicity of *Haemophilus influenzae*. *Journal of Medical Microbiology*, 18: 1-16.

Urwin, G, JM Musser and MF Yuan. 1995. Clonal analysis of *Haemophilus influenzae* type b isolates in the United Kingdom. *Journal of Medical Microbiology*, 43: 45-49.

van Alphen, L, L Geelen, K Jonsdottir, AK Takala, H Kayhty and HC Zanen. 1987. Distinct geographic distribution of subtypes of *Haemophilus influenzae* type b in Western Europe. *Journal of Infectious Diseases*, 156: 216-218.

van Alphen, L, T Riemens, J Poolman, C Hopman and HC Zanen. 1983. Homogeneity of cell envelope protein subtypes, lipopolysaccharide serotypes, and biotypes among *Haemophilus influenzae* type b from patients with meningitis in the Netherlands. *Journal of Infectious Diseases*, 148: 75-81.

van Alphen, L, L Spanjaard, A Vanderende and J Dankert. 1994. Predicted disappearance of *Haemophilus influenzae* type b meningitis in the Netherlands. *Lancet*, 344: 195.

van Alphen, L, AK Takala, L Geelen-van den Broek, J Dankert and J Eskola. 1992. Changes in the distribution of *Haemophilus influenzae* type b clones associated with widespread infant vaccination in Finland. *Journal of Infectious Diseases*, 166: 1340-1345.

van Belkum, A, B Duim, A Regelink, L Moller, W Quint and L van Alphen. 1994. Genomic DNA fingerprinting of clinical *Haemophilus influenzae* isolates by polymerase chain reaction amplification: comparison with major outer membrane protein and restriction fragment length polymorphism analysis. *Journal of Medical Microbiology*, 41: 63-68.

van Belkum, A, M Struelens, A de Visser, H Verbrugh and M Tibayrenc. 2001. Role of genomic typing in taxonomy, evolutionary genetics, and microbial epidemiology [Review]. *Clinical Microbiology Reviews*, 14.

van Ham, SM, L van Alphen, FR Mool and JPM van Putten. 1994. The fimbrial gene cluster of *Haemophilus influenzae* type b. *Molecular Microbiology*, 13: 673-684.

van Loo, IHM, KJ Heuvelman, AJ King and FR Mooi. 1994. Multilocus sequence typing of *Bordetella pertussis* based on surface protein genes. *Journal of Clinical Microbiology*, 40: 1994-2001.

Vitovski, S, KT Dunkin, AJ Howard and JR Sayers. 2002. Nontypeable *Haemophilus influenzae* in carriage and disease: a difference in IgA1 protease activity levels. *Journal of the American Medical Association*, 287: 1699-1705.

Ward, JI, MKW Lum and DB Hall. 1986. Invasive *Haemophilus influenzae* type b disease in Alaska: background epidemiology for a vaccine efficacy trial. *Journal of Infectious Diseases*, 153: 17-26.

Ward, JI, HS Margolis, MKW Lum, DW Fraser and TR Bender. 1981. *Haemophilus influenzae* disease in Alaskan Eskimos: characteristics of a population with an unusual incidence of invasive disease. *Lancet*, i: 1281-1284.

Wayne, LG, DJ Brenner, RR Colwell, PAD Grimont, O Kandler, MI Krichevsky, LH Moore, WEC Moore, RGE Murray, E Stackebrandt, MP Starr and HG Truper. 1987. Report of the *ad hoc* committee on the reconciliation of approaches to bacterial systematics. *International Journal of Systematic Bacteriology*, 37: 463-464.

Webb-Wagg, KD. 1997. Honours thesis. The genetic diversity of nontypeable *Haemophilus influenzae*. University of Canberra. Australia.

Weiller, G and AJ Gibbs. 1993. DIPLOMO: Distance Plot Monitor, version 1.3. Australian National University. http://life.anu.edu.au.

Weinberg, GA, A Ghafoor, Z Ishaq, NK Nomani, M Kabeer, F Anwar, MI Burney, W Qureshi, JM Musser, RK Selander and DM Granoff. 1989. Clonal analysis of *Haemophilus influenzae* isolated from children from Pakistan with lower respiratory tract infections. *Journal of Infectious Diseases*, 160: 634-642.

Weiser, JN. 1993. Relationship between colony morphology and the life cycle of *Haemophilus influenzae*: the contribution of lipopolysaccharide phase variation to pathogenesis. *Journal of Infectious Diseases*, 168: 672-680.

Weiser, JN, A Williams and ER Moxon. 1990. Phase-variable lipopolysaccharide structures enhance the invasive capacity of *Haemophilus influenzae*. *Infection and Immunity*, 58: 3455-3457.

Weiss, DPL, P Coplan and H Guess. 2001. Epidemiology of bacterial meningitis among children in Brazil, 1997-1998 [Spanish: English abstract]. *Revista de Saude Publica*, 35: 249-255.

Whittam, TS, H Ochman and R Selander. 1983. Multilocus genetic structure in natural populations of *Escherichia coli*. *Proceedings of the National Academy of Sciences USA*, 80: 1751-1755.

Willett, HP. 1988. *Haemophilus*. *In* Zinsser Microbiology. Joklik, WK, HP Willett, DB Amos and CM Wilfert. Appleton & Lange, East Norwalk: 393-401.

Woese, C. 1987. Bacterial evolution. *Microbiological Reviews*, 51: 221-271.

Yakubu, DE, JRA Fariborz and TH Pennington. 1999. Molecular typing methods for *Neisseria meningitidis*. *Journal of Medical Microbiology*, 48: 1055-1064.

Yang, YH, XZ Shen, ZF Jiang, XY Liu, ZQ Leng, DL Lu, JT Rao, JQ Liu and LF Chang. 1998. Study on *Haemophilus influenzae* type b diseases in China - the past, present and future. *Pediatric Infectious Disease Journal*, 17: S 159-S 165.

Young, JPW. 1989. The Population Genetics of Bacteria. *In* Genetics of Bacterial Diversity. Hopwood and Chester, KE. Academic Press: 417-438.

Zwahlen, A, JS Kroll, LG Rubin and ER Moxon. 1989. The molecular basis of pathogenicity in *Haemophilus influenzae*: comparative virulence of genetically related capsular transformants and correlation with changes at the capsulation locus *cap*. *Microbial Pathogenesis*, 7: 225-235.

Zwahlen, A, LG Rubin and ER Moxon. 1986. Contribution of lipopolysaccharide to pathogenicity of *Haemophilus influenzae*: comparative virulence of genetically related strains in rats. *Microbial Pathogenesis*, 1: 465-473.